Title
Chemical Tools for Investigating Galectin Dynamics

Author
Belardi, Brian

Publication Date
2014

Peer reviewed|Thesis/dissertation

Chemical Tools for Investigating Galectin Dynamics

By

Brian Daniel Belardi

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Carolyn R. Bertozzi, Chair
Professor Matthew B. Francis
Professor Jay T. Groves
Professor Daniel A. Fletcher

Fall 2014

Chemical Tools for Investigating Galectin Dynamics

By Brian Daniel Belardi

Abstract

Chemical Tools for Investigating Galectin Dynamics

by

Brian Daniel Belardi

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Carolyn R. Bertozzi, Chair

Pattern recognition is an essential feature governing many biological processes, including metazoan development, infection, and tumorigenesis. Glycans form a unique pattern on the surfaces of cells. There, these branched, intricate structures are poised to direct communication between neighboring cells and tissues. How cells interpret the staggering complexity associated with the glycome, the repertoire of glycan structures expressed in cells and tissues, though is still not well understood. Galectins comprise one family of glycan-binding proteins that are thought to play a role in decoding glycomic patterns. They have been implicated in cell adhesion and signaling, proliferation, membrane organization, and transcript processing. Yet, the mechanistic underpinnings of the pleiotropic functions of galectins remain obscure due to challenges associated with studying lectin-glycan interactions.

Monitoring and modulating glycan patterns present on different cell surface glycoconjugates would greatly facilitate our understanding of galectin-dependent phenomena. Unfortunately, glycosylation is a post-translational modification that is dependent on the collective action of hundreds of enzymes. The final glycans appended to proteins and lipids are not directly encoded in the genome, and as such, there is no obvious means to tag or tailor complex, galectin ligands *in vivo*. This situation is further exacerbated by the properties of galectins themselves. Galectin-1, for instance, self-assembles *in vitro* and traverses cell membranes, occupying extracellular, cytosolic, and nuclear spaces. Therefore, isolating galectins' dynamic organization in each subcellular context is imperative, yet non-trivial experimentally.

This dissertation describes the development of new chemical and biophysical tools for studying galectin dynamics and their contribution to cellular physiology and pathology. Chapter 1 explores and summarizes recent advances in chemical methodology that report on mammalian lectin function *in vivo*. In Chapters 2 and 3, I present a platform for monitoring galectin-mediated cross-linking on living cells with synthetic glycoprotein mimics. Chapter 4 outlines a technique to image galectin ligands, i.e. specific protein glycoforms, in cultured cells and tissue slices. Finally, Chapter 5 is dedicated to clarifying galectin's influence on mammary gland morphogenesis and breast cancer progression. Therein, I discuss harnessing the tools described in Chapters 1-4 to reveal an unprecedented, galectin-based mode of information transfer between

extracellular glycomic signatures and nuclear transcription machinery that drives mammary epithelial invasion.

This dissertation is dedicated to my Java Runtime Environment (J.R.E.)
and my family

Chemical Tools for Investigating Galectin Dynamics

Table of Contents

List of Figures

List of Tables

List of Schemes

Acknowledgements

Graduate school is transitory by nature. Individuals fly in and out of your purview as you advance through a shifting, yet finite landscape. However, a number of friends, collaborators, and mentors permanently shape your experience and send ripples that resonate with you for years to come. I would like to thank these people for making graduate school such a rewarding and wonderful time. First and foremost, I thank my advisor, Carolyn Bertozzi. She has molded me into the scientist that I am today. Carolyn has supported my ideas, personality, and development with unwavering enthusiasm during my graduate school tenure. Often my scientific predilections led to research pursuits outside her area of expertise. She nevertheless offered the freedom and encouragement to follow directions I found exciting and unexplored. Carolyn is a brilliant scientist, writer, and communicator, and I have learned a great deal from her generous advice and wisdom over the years. Perhaps most importantly, Carolyn has fostered a passion for viewing science with awe and amusement, and I am truly grateful for that.

I would like to thank members of the Bertozzi laboratory for creating such a stimulating and conducive research environment. My early mentors included John Jewett, Mike Boyce, and Ellen Sletten, who were patient with a young, confident but naïve scientist, and I am forever indebted to them for their instruction and guidance during those formative years. Some of the most memorable moments in lab were spent with fellow inhabitants of room 802. I would first like to thank Karen Dehnert. Karen was a friend and a mentor. She was funny and incredibly intelligent, and I appreciated her gentle direction and perspective. Throughout graduate school, I have relied on Jason Hudak. Jason and I formed a strong, lasting bond during my first year. And I had the pleasure of watching Jason develop into an innovative scientist and a respected figure in the field. I am grateful for all the assistance and encouragement he provided and am lucky to call him a friend and collaborator. Peyton Shieh and I have shared countless laughs over the years. Peyton is a towering intellect, and I am thankful for his thoughtful suggestions. He has consistently been a helpful sounding board for new ideas, and I look forward to what he accomplishes next. More recently, David Spiciarich has been a wonderful collaborator, and I thank him for being a caring researcher and a beacon of light within the lab. Although Christina Woo and I only overlapped briefly, she has graciously assisted with my transition from graduate student to post-doc. I am confident that she will be a successful scientist and academic for many years to come. I also must thank Adam de la Zerda, who took the time to teach me how to approach instrumentation from a physical perspective. Brendan Beahm, Lauren Wagner, Ben Swarts, Mason Appel, and Elliot Woods have all been stalwart supporters of my work and dear friends. They have brought a healthy amount of levity to lab, and I am grateful for that. I would also like to express my sincere gratitude to Asia Avelino and Olga Martinez for bending rules and manipulating situations to advance my project and career.

I would be absolutely remiss if I did not acknowledge my classmates and closest friends in graduate school, Aaron Lackner, Geoffrey O'Donoghue, Jeffrey Wu, Matthew Ramirez, and Kristen Seim. This group consisted of some of the most gifted, compassionate, and funny individuals I have encountered in my life. The ease with which they transitioned between intense intellectual arguments and commonplace banter was remarkable. They propped me up innumerable times and elevated my sense of self. I am certain that they will be lifelong friends and confidants, and I thank them for enriching my graduate school experience.

Several collaborators have directly contributed to work presented in this dissertation. I would first like to thank Jay Groves. While others doubted my atypical research approaches and interests, Jay cultivated them and provided mentorship in the early stages of graduate school. He welcomed me into his lab, and I am grateful for his constant support. Adam Smith was an outstanding teacher and advocate, and I thank him for his help. I also would like to single out my good friend and collaborator, Ramray Bhat. Ramray and I have a kinship with one another and share surprisingly similar worldviews. To me, though, Ramray is an inspirational figure. He pushed me to love the essence of science and think creatively at every turn. My time with him was full of philosophical musings, smiles and excitement. I can't thank him enough for all of his encouragement. I'm grateful to Mina Bissell who offered enthusiastic support for my work and provides an example of how to age gracefully. I also thank Holly Aaron for her expert technical assistance and positive demeanor during many long days and nights.

I worked with two immensely talented undergraduates at Berkeley. First, I trained and advised Yichao Huang from Tsinghua University. He was a pleasure to be around and an angelic presence in my life. I also want to thank Angelo Solania for his time in the lab. Angelo was an impressive scientist and exuded maturity for someone his age. I know both will go on to have long, successful careers in science.

Finally, I'd like to close by acknowledging Julia Ende and my family for their endless support and their steadfast faith in me during these years at Berkeley. I am truly thankful for having my love, Julia, by my side throughout graduate school. She is a special person on this earth, a brilliant woman with an inexplicable gift to make me laugh, feel alive, comforted, and confident, all at once. It has been a blessing to grow intellectually together. My family - my mom, my dad, my sister, and my brother - have made numerous sacrifices for me during my tenure at Berkeley. They were always interested in my work; my mom probably understood my research better than most members of my lab. And they understood what it meant to be a graduate student. Their support was absolutely instrumental in every one of my accomplishments. I can't thank them enough for this privileged experience. This dissertation is dedicated to them.

Chapter 1

Chemical Methods for Probing the Dynamics of Mammalian Lectins *In Vivo*

Chapter 1. Chemical Methods for Probing the Dynamics of Mammalian Lectins *In Vivo*

Introduction

The glycan appendages to proteins and lipids, collectively termed the glycome, are an information-rich set of molecular cues that influences almost every facet of molecular, cellular, and organismal biology, including, but not limited to, protein stability and function, cellular trafficking, signal transduction, cell-cell recognition, development, host defense and progression of pathology.[1] The glycome exceeds the combinatorial complexity of other classes of biomolecules, e.g. proteins and nucleic acids.[2] This diversity is due to the action of ~250 glycosyltransferases that catalyze the addition of different glycan moieties onto their substrates with varying regio- and stereo-chemical specificities. While physical properties of these modifications can partially account for their biological effect, the overwhelming number of diverse glycan elaborations represents a molecular code that is interpreted by the local environment. In mammals, a class of glycan-binding proteins, lectins, is responsible for deciphering this glycan code.[3,4]

Lectin research during the past century has made significant strides towards the isolation of different lectin family members and the understanding of lectin binding preferences, hemagglutination activity and structure *in vitro*.[5] This pursuit has been extremely productive as lectins are now routinely used as biological reagents in glycoprotein purification and cell staining as well as in biomedical-related applications. However, unraveling the *in vivo* glycan partners and functions of many mammalian lectins has been hampered by several features intrinsic to lectin-glycan interactions. First, studying the glycan partners for an individual lectin is exceptionally challenging using traditional tools. These ligands comprise a heterogeneous set of glycoconjugates with similar glycan compositions but disparate underlying protein and lipid scaffolds. Glycosylation, a post-translational modification, is not amenable to genetic manipulations, and as such, encoding a specific glycan structure with genetic engineering is not possible. As well, not all *in vitro* glycan ligands translate to an *in vivo* setting.[6] In the context of a cell surface, it is probably only a subset of possible glycoprotein ligands, for instance, that actually interacts with a lectin-of-interest. The protein backbone is known to complicate glycan presentation, sometimes restricting glycan conformation and localization and even modulating affinity through cooperative interactions.[7,8] Further still, lectins have convergently evolved to possess weak affinity for their monomeric glycan ligands (K_d's of 100 uM to 1 mM). Multivalency, the presence of numerous glycan ligands on a single target, is responsible for the increased avidity of lectins for their targets.[9,10] However, *in vivo*, these multiple glycan contacts may be heterogeneous and thus difficult to fully elucidate.

Another overlooked, but nonetheless vital feature, of lectin-glycan interactions relates to lectin dynamics. Early notions of mammalian lectin function dealt with glycan-binding proteins as docking sites for other cells and soluble glycoconjugates, i.e. sources of adhesion.[11] Many lectins have the capacity to self-assemble though, either in solution or when embedded in the membrane, begging the question of what effect this ability has on glycan interactions.[12] Aside from an increase in binding avidity, oligomeric lectins resemble adaptor proteins and can act as an organizing principle *in vivo*. They can cluster their glycoconjugate ligands into functional domains or increase the probability that protein-protein associations transpire. As well, it is now clear that kinetic rate constants (k_{on} and k_{off}) for this assembly process vary under different

glycosylation and cellular conditions, tuning the influence of lectin multimerization.[13] Since most lectin families are thought to lack regulatory domains or allosteric control, lectin trafficking and sub-cellular localization, and their respective dynamics, control how and when lectins exert their function. Unfortunately, traditional biological tools are not able to follow many of these time-dependent and sub-diffraction events *in vivo*. For example, tagging lectin monomers, which can be as small as ~15 kDa,[14] with fluorescent proteins (molecular weights of ~25 kDa)[15] can perturb lectin self-assembly, and visualizing this assembly with conventional microscopy obscures much of a lectin's organizational properties.

Rapid advances in chemical methods have begun to meet the many challenges associated with unveiling lectin's contributions *in vivo*. In this chapter, the focus will be on recent contributions that attempt to answer either 1) what native glycan structures mammalian lectins recognize or 2) what the dynamics of different lectins are and how they relate to lectin function. The various chemical strategies can be divided into three categories: metabolic cross-linking sugars, glycoprotein mimics, and site-selective tagging and imaging technologies (Figure 1.1). While this is by no means an exhaustive list of techniques for observing lectin binding and dynamics, these tools have transformed our understanding of lectin-glycan interactions *in vivo* and will continue to reveal unexpected features of mammalian lectins in the future.

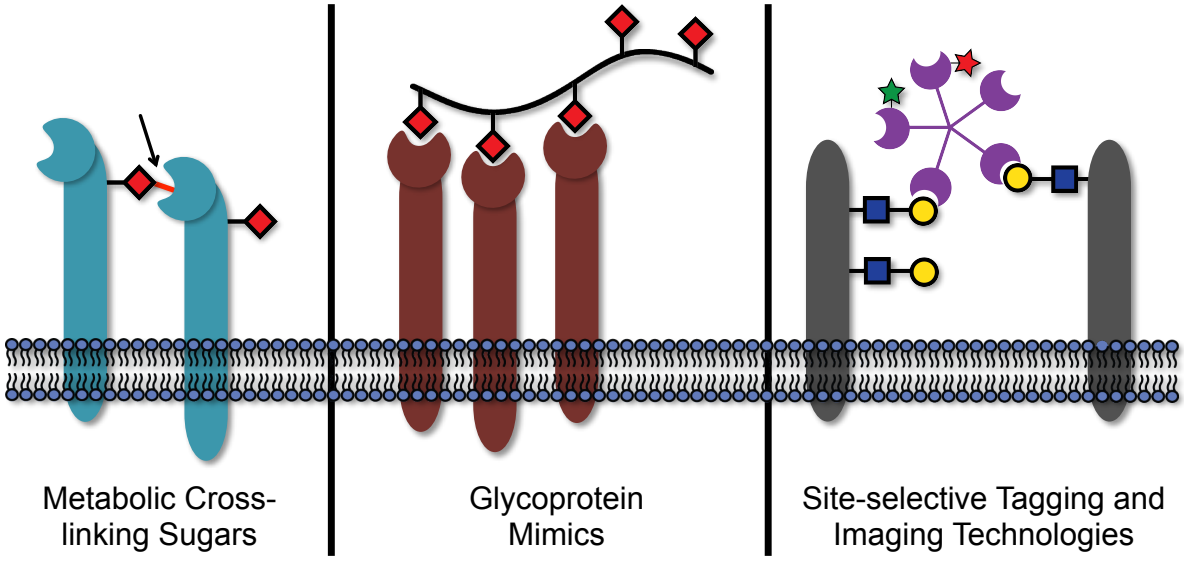


Figure 1.1. Recent chemical methods have begun to elucidate the binding partners and dynamic behavior of mammalian lectins *in vivo*. Metabolic cross-linking sugars can report on the *in situ* interactions between lectins and glycans by incorporating into endogenous cell surface glycoproteins. Glycoprotein mimics, on the other hand, are synthetic, multimeric structures capable of revealing the properties and biological consequences of lectin organization. New tagging and imaging technologies aim to complement these glycan-based strategies and promise to monitor lectin trafficking and organization *in vivo*.

Metabolic Cross-linking Sugars

A single lectin can associate with glycans derived from a variety of linkages, such as N-linked and O-linked glycoproteins as well as glycolipids. N-linked glycosylation occurs co-translationally or post-translationally in the endoplasmic reticulum (ER).[1] The addition of a tetradecasaccharide to asparagine residues within the sequon, NXS/T, is accomplished through the action of a single enzyme complex, oligosaccharyltransferase (OST). Mucin-like O-linked glycans, on the other hand, are initiated post-translationally in the Golgi apparatus, where a member of the polypeptide N-acetylgalactosaminyltransferase (ppGalNAcTs) family modifies a serine or threonine residue with the monosaccharide, N-acetyl galactosamine (GalNAc).[1] The early steps in these glycosylation pathways are quite distinct, generating unique glycan structures near the peptide backbone. However, as both types of glycoproteins in addition to glycospingolipids transit through the Golgi, later glycosyltransferases can act on each class of glycans and produce terminal glycosylation patterns that are similar or identical. Metabolic oligosaccharide engineering provides a means to target similar glycan structures associated with diverse scaffolds.

Metabolic engineering of glycosylation, first demonstrated by Reutter and co-workers,[16,17] relies on the native cellular biomachinery to incorporate non-natural glycans into native glycoconjugates. By extending the acyl side chain of N-acetylmannosamine (ManNAc), which is part of the salvage pathway of sialic acid, nonphysiological versions of sialic acid were shown to replace the native monomer *in vivo*.[18] This idea was greatly expanded by Bertozzi and co-workers in a series of experiments.[19,20] Chemical derivatives of ManNAc and sialic acid, bearing bioorthogonal functionalities, were found to be tolerated by the sialic acid biosynthetic enzymes, the CMP-sialic acid Golgi transporter, as well as by a number of sialyltransferases, to afford derivatives of cell surface sialoside structures. As well, the Bertozzi lab and others have incorporated unnatural reporter groups into glycans containing GalNAc,[21] N-acetylglucosamine (GlcNAc),[22] fucose,[23] and xylose residues[24]. These pioneering efforts provided evidence that similar glycan structures on a variety of different protein and lipid scaffolds could be targeted using a metabolic-based approach. Recently, a number of reports have detailed the combination of metabolic labeling, where reporter groups attached to cell surface glycans are probed with fluorescent conjugates of bioorthogonal reaction partners, and Förster Resonance Energy Transfer (FRET) for imaging the precise location of specific glycoprotein ligands.[25–27] These methods, while certainly expanding the scope of locating glycoconjugate ligands *in vivo*, are not necessarily reporting on *bona fide* lectin-glycan interactions. Binding events between lectins and glycans are plagued by transient non-covalent interactions with high dissociation constants and are thus not easily isolable. If possible, taking a snapshot of cellular interactions would provide a means to discriminate real from theoretical binding partners *in vivo*.

Carbenes and similar electron-deficient species, highly reactive functionalities, are known to undergo C-H and heteroatom-H insertion reactions, forming stable covalent adducts rapidly.[28] Fast kinetics associated with these insertion reactions can be exploited to trap short-lived binding events in biological samples. Covalent cross-linking has recently been utilized to identify lectin-glycan interactions *in vivo*.

Siglec-2, or CD22, is a member of the sialic acid-binding immunoglobin-like lectin (siglec) family and modulates immune cell activation.[29] The sialic acid binding domain of CD22, an integral membrane protein, is located in the extracellular space. The other functional domain of CD22 is cytosolic and has an immunoreceptor tyrosine-based inhibitory motif (ITIM) that

when phosphorylated recruits SH2 domain-containing phosphatase 1.[30] This response inhibits downstream signaling associated with B-cell receptor (BCR) triggering and establishes a threshold for activation, which limits autoimmunity. On *resting* B cells, however, it was hypothesized that *cis* interactions with other cell surface glycoproteins masked CD22's lectin domain, preventing CD22 and BCR co-localization. To determine the precise *cis* interactions of CD22, Han *et al.* derivatized the C-9 position of sialic acid with an azidoaryl species (9-AAz-NeuAc).[31] Azidoarenes form nitrenes after UV light exposure and subsequent expulsion of nitrogen and can capture proximal proteins by C-H insertion. To identify the exact glycoprotein partners *in vivo*, BJAB K20 cells, a B cell line deficient in sialic acid, were fed 9-AAz-NeuAc. After establishing that the azido functionality survived metabolic incorporation by performing a Staudinger ligation between cell surface azides and a phosphine-biotin probe, cells were first treated with 9-AAz-NeuAc and then irradiated with UV light to induce cross-linking. After immunoprecipitation of CD22 from cell lysates, different antibodies were used to test for glycoproteins that mask CD22 on B cell surfaces. Interestingly, the only glycoprotein found to cross-link CD22 was other CD22 molecules. Glycoproteins that bound soluble CD22 fusion constructs *in vitro*, for example CD45 and CD19, were not cross-linked to CD22 *in vivo*, emphasizing the power of this metabolic-based approach. In addition, the cross-links were found be high molecular weight (>400 kDa) species, suggesting that CD22 associates with itself and forms multimers on the cell surface. This last finding also reflects the unique dynamics of CD22 *in vivo*. CD22's cytosolic domain has a motif that is recognized by a component of clathrin-coated pits and therefore may reinforce CD22 self-association and microdomain formation on resting B cells.

Paulson and co-workers wanted to also apply this same photocross-linking strategy to the *trans* ligands of CD22.[32] During B cell engagement, glycoproteins on apposing cells unmask CD22, changing its dynamics. Subsequent localization to regions enriched in BCR downregulate immune signaling. The authors generated a list of potential glycoprotein candidates that interact during B cell-B cell contact by first incubating K20 cells with 9-AAz-NeuAc in serum-free media. Intact cells were then treated with the soluble CD22-Fc chimera, a surrogate for B cells, and exposed to UV light. After lysis, cross-linked species were isolated by a Protein G-resin and separated by gel electrophoresis. In-gel trypsin digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS) identified 24 proteins, whose peptides had high spectral counts compared to non-UV radiation exposed cells. To corroborate their *trans* ligand list as well as increase the number of identified candidates through stable isotope labeling by amino acids in cell culture (SILAC), 9-AAz-NeuAc positive cells were grown either in media containing light ([^{12}C]lysine and [^{12}C,^{14}N]arginine) or heavy ([^{13}C]lysine and [^{12}C,^{15}N]arginine) amino acid isotopes. Both populations were separately incubated with CD22-Fc and only the heavy isotope-fed cells were irradiated. After combining both populations, purifying CD22 cross-linked species, and performing an on-bead trypsin digest, multi-dimensional protein identification technology-tandem mass spectrometry (MuDPIT-MS/MS) was used to discriminate and quantify proteins that were more abundant after photocross-linking. This complementary method yielded 19 proteins, all of which were annotated as glycoproteins, and when combined with the previous list, resulted in 27 possible glycoprotein *trans* ligands of CD22. The authors proceeded to examine each of these candidates in a cell-cell contact photocross-linking experiment. K20 cells were labeled with 9-AAz-NeuAc and overlaid onto Chinese Hamster Ovarian (CHO) cells expressing V5-tagged CD22. After UV exposure, lysates were formed and CD22 was immunoprecipitated with α-V5. After probing the cross-linked species for the presence of each

of glycoprotein candidate, three glycoproteins were found to be interacting with CD22 in *trans*: IgM, CD45, and Basigin. Of these, IgM was the only ligand that showed robust cross-linking. In a series of knockdown studies, the authors provided striking evidence that both CD22 and IgM dynamics are altered during B cell-B cell contact as the partners redistribute and localize preferentially to the sites of cell contact in a lectin- and glycan-dependent manner. Again, this work emphasizes how both glycan structure and dynamics are intimately involved in dictating binding partners *in vivo.*

There are limitations associated with functionalizing C-9 of sialic acid. This position can undergo further modifications, such as 9-*O*-acetylation, 9-*O*-lactylation, and 9-phosphorylation,[33] which precludes 9-AAz-NeuAc's use in situations where elaborated versions of sialic acid are under investigation. Tanaka *et al.* circumvented this problem by replacing the acetyl group of ManNAc with an N-acyl diazirine photo-labile functionality.[34] Diazirines are known to generate carbenes after UV illumination and therefore cross-link binding partners that are in close association. Another liability with 9-AAz-NeuAc is its low incorporation levels in wildtype cell lines, but a smaller structural perturbation, such as a diazirine unit, should compete with the endogenous pool of sialic acid. In addition to K20 cells, wildtype K88 cells fed the fully protected diazirine-ManNAc analogue, Ac_4ManNDAz, incorporated SiaDAz into cell surfaces glycans, and once there, SiaDAz cross-linked siglec CD22. This result provided further evidence that CD22 appears to form *cis* multimers on *wildtype* resting B cells. Kohler and coworkers went on to show that the chain length of N-acyl diazirines was crucial for proper incorporation.[35] ManNDAz derivates containing two (2me), three (3me), or four (4me) methylene chain units were synthesized and evaluated for replacement of endogenous sialic acid and subsequent cross-linking capability (Figure 1.2). In Daudi B cells, only Ac_4ManNDAz (2me) was found to cross-link CD22, which defines the maximal structural perturbation permissible during metabolic transformation. This is just one example of how the ability to construct molecules can pinpoint the necessary properties for probing lectin-glycan interactions.

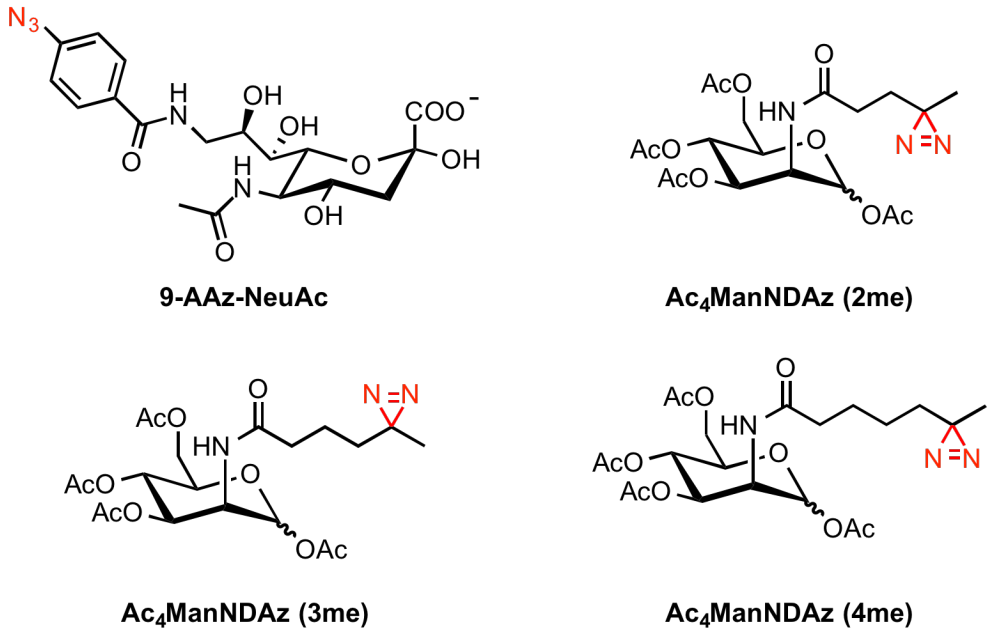


Figure 1.2. Structures of metabolic cross-linking sugars that upon UV exposure generate nitrenes or carbenes (red). These intermediates then react with lectins *in vivo* to form stable covalent adducts.

In less than a decade, metabolic cross-linking sugars have unearthed surprising features of CD22 on B cells. Many glycoproteins are modified with an N-acetylneuraminic acid α(2,6) galactose structure (Neu5Ac-α(2,6)-Gal), but among the plethora of possible binding partners of the lectin CD22, only CD22 and IgM appear to govern *cis* and *trans* interactions *in vivo,* respectively. The Kohler lab has recently expanded the library of cross-linking sugars to other glycan monomers,[36] and we anticipate that many of these derivatives will be used to detect additional mammalian lectin-glycan interactions in the future. We also envision research towards inverting the current placement of the cross-linking functional group. If the cross-linking moiety is placed within the lectin, then this may provide a means to efficiently capture *in vivo* glycoprotein ligands, particularly in cases where the carbohydrate specificity of a lectin is not well understood. Incorporation of photoreactive benzophenones into proteins has already been accomplished in mammalian cells through unnatural amino acid incorporation.[37] This technology could easily be expanded to the recognition domain of most lectins *in vivo*.

Glycoprotein Mimics

Many glycoproteins, especially mucin-type O-linked glycoproteins, are heavily elaborated biomolecules with repeating motifs of glycosylation. The glycan signatures displayed across a single protein backbone can vary dramatically, making analysis of lectin-glycoprotein function difficult. Chemists have harnessed biomimicry for decades in order to construct molecular surrogates that emulate the essential features of biomolecules but differ in tunable composition and properties.[38] Three general classes of glycoprotein mimics have been successfully synthesized and used to bind mammalian lectins both *in vitro* and *in vivo*: glycopolymers,[39] linear polymers with pendant glycans; glycodendrimers,[40] highly branched

architectures with terminal glycan residues; and glycoparticles,[41] cross-linked organic polymers or spherical inorganic materials that are functionalized with glycans through various surface chemistries. As opposed to the native heterogeneous glycoforms of proteins, these synthetic glycoconjugates offer the ability to incorporate homogeneous glycans structures into a repeating sequence. This section will focus on glycopolymers as mimics of glycoproteins, since recent synthetic methods, offering unparalleled control over their construction, have yielded profound insights into lectin biology.

The original synthesis of glycopolymers dates back to the 1970's.[42] Since then, glycopolymers have found widespread use as inhibitors, for instance during bacterial infection[43], and as tools for describing the binding modes of multivalency *in vitro*.[44–46] In the past decade, chemical biologists have continued to push glycopolymers into the biological milieu.[47] Free glycopolymers resemble soluble, secreted glycoproteins, and owing to improved synthetic methodology, it is now possible to prepare glycopolymers that mimic membrane-bound glycoproteins. With these tools in hand, a number of questions relating to lectin dynamics are now accessible and have been addressed for three large families of lectins, namely C-type lectins, siglecs, and galectins.

Selectins, calcium-dependent lectins (C-type lectins), are critical mediators of cell adhesion, especially of immune cell homing.[48] The three selectins, P, L, and E- selectins, are expressed on different cell types, platelets, leukocytes, and endothelium, respectively. During the inflammatory response, leukocytes roll along endothelial cells, adhere to the vascular substrate and extravasate into the damaged tissue. As a corollary, the overexpression of L-selectin can lead to enhanced leukocyte migration and inflammatory damage, and thus L-selectin density is tightly regulated on leukocyte surfaces. The interaction between L-selectin and GlyCAM-1, an O-linked glycoprotein, plays a major role in leukocyte rolling along the endothelial wall. Since GlyCAM-1 possesses sulfated versions of the sialyl Lewis x epitope (sLe^x), Neu5Ac-α2-3Gal-β1-4(Fuc-α 1-3)GlcNAc, Mowery *et al.* were interested in how the differential sulfation of this epitope affected L-selectin signaling and regulation.[49] The authors employed a type of living polymerization, ring opening metathesis polymerization (ROMP), to produce well-defined glycoprotein mimics. Living polymerization is a chain-growth method that suppresses the rate of termination compared to the rate of propagation, thereby rendering macromolecules with precise degrees of polymerization (DOP) and polydispersities (PDI).[50] ROMP involves strained cyclic olefins and the use of transition metal complexes, containing ruthenium or molybdenum, to mediate their polymerization.[51] Starting from functionalized norbornenes, different glycopolymers bearing sialyl Lewis x or mono- and disulfated analogs of galactose, Lewis x, and sialyl Lewis x were prepared (DOP = 15). Acting as soluble glycoprotein mimics, the glycopolymers were tested for their ability to downregulate L-selectin on lymphoid cells. The 3',6-disulfo Lewis x glycopolymer caused the most significant decrease in surface L-selectin levels after binding. Consistent with these results, glycopolymer treatment resulted in soluble L-selectin in the supernatant, therefore indicating that shedding, or proteolytic release, of L-selectin was responsible for the lectin's downregulation. None of the monovalent glycan equivalents promoted L-selectin shedding. The data together provide evidence that clustering of L-selectin by disulfated glycoprotein mimics on the cell surface signals it's proteolytic release. It also appears that sulfation and therefore Coulombic forces as opposed to sialylation of glycoprotein ligands are more important factors in L-selectin shedding *in vivo*.

To address how spacing of these glycan epitopes along a glycoprotein might modulate the signaling of L-selectin clusters, well-defined glycoconjugates were formed using alanine-rich

recombinant polypeptides.[52] Glycan conjugation to glutamic acid residues at defined positions yielded well-defined, α-helical glycoprotein mimics. Specifically, disulfated galactosyl amines were coupled to the polypeptide backbone with HBTU and arranged at two different distances, 1.7-3.5 nm and 3.5-5.0 nm. Jurkat T cells were incubated with these glycopolypeptides, and the amount of L-selectin shedding was determined by an enzyme-linked immunosorbent assay (ELISA) for the protein. As anticipated, monovalent glycans and an unstructured glycopolypeptide did not induce L-selectin shedding. However, while the 1.7-3.5 nm construct resulted in L-selectin cleavage, the 3.5-5.0 nm glycopolypeptide, unexpectedly, did not. Once again, this discrepancy in shedding highlights how the dynamics of selectin organization on cell surfaces has functional consequences. Formation of L-selectin clusters are thought to lead to calmodulin (CaM) release by the cytoplasmic tail of the protein, which is then followed by a conformational change that causes the protein to be susceptible to sheddase activity. Results from glycoprotein mimics outlined above add another parameter that influences signaling and shedding of L-selectin. L-selectin monomers must be within sufficient proximity to one another after binding for signaling to occur, a spatial requirement that appears necessary for subsequent shedding.

Dendritic cell-specific ICAM-3-grabbing nonintegrin (DC-SIGN) is a member of the C-type lectin class and carries out a multitude of functions on DCs through interactions with high mannose glycans and fucosylated Lewis-type antigens on ICAM-3 and ICAM-2.[53] DC-SIGN stabilizes immunological synapases during T-cell engagement and can mediate trans-endothelial migration. It was also thought to be involved in antigen uptake and processing. A glycomimetic was subsequently developed that could undergo DC-SIGN-mediated uptake.[54] Bovine serum albumin (BSA), acting as a polymeric backbone, was decorated with mannose mimetics bearing either native mannosyl or fucosyl residues using lysine chemistry (~16 copies per BSA). Fluorophore-functionalized BSA glycoconjugates were administered to Raji cells, a B cell line, stably transfected with the gene for DC-SIGN (Raji/DC-SIGN). The Raji/DC-SIGN cells, but not the *wt* Raji cell line, internalized all three glycoconjugates, including the unnatural mannose mimetic-BSA. The glycoprotein surrogates also stimulated JNK signaling in response to interaction with DC-SIGN. Phosphorylation of JNK was not seen in other studies of DC-SIGN binding, leading to the proposal that DC-SIGN reorganization and perhaps clustering is necessary for signaling to take place. C-type lectin-dependent uptake may be a general feature in antigen processing as polyacrylamide (PAA)-GalNAc glycopolymers are readily internalized by DCs through Macrophage Galactose-type C-type Lectin 2 (MGL2), resulting in their presentation in DCs.[55]

As discussed above, CD22 belongs to the siglec family of lectins. An important siglec on B cells, CD22 downregulates immune cell activation via recruitment of SHP-1 phosphatase to the site of receptor signaling. *Cis* α2,6-linked sialic acid-CD22 interactions were shown to mask CD22 on resting B cells. But one question remained: what are the properties of *trans* ligands that promote escape from masked CD22 multimers? Collins *et al.* examined this question by constructing a series of sialoside glycoprotein mimics.[56] Two different molecular mass biotinylated PAA polymers, 30 kDa and 1000 kDa, were used as starting materials for the attachment of sialosides, including NeuAc and its C-9-biphenyl derivative (BPC-NeuAc) as well as the native N-glycolylneuraminic acid (NeuGc) and its C-9-biphenyl derivative (BPA-NeuGc). Aromatic groups at the C-9 position of sialic acid had been previously shown to increase affinity toward CD22.[57] The prepared glycopolymers were added to either BJAB or Daudi B cells, and

flow cytometry was used to detect glycopolymer binding. Only the 1000 kDa BPC-NeuAc-PAA and BPA-NeuGc-PAA polymers adhered to the B cell surface. The native sialoside polymers, NeuAc-PAA and NeuGc-PAA, in addition to the 30 kDa glycopolymers did not exhibit cell surface retention. A simple interpretation of these findings is that CD22 exists in a dynamic 'equilibrium' on the surface of B cells. If presented with ligands possessing a sufficient threshold of affinity, the CD22 binding 'equilibrium' favors unmasking, whereas if the glyconconjugate ligands are too weak, then the CD22 'equilibrium' favors self-association. Only in the case when CD22 is presented with highly multivalent glycopolymers, 1000 kDa, containing high affinity glycan derivatives, i.e. BPC-NeuAc and BPA-NeuGc, does unmasking occur. The authors go on to provide evidence that the CD22-bound PAA glycopolymers undergo clathrin-mediated endocytosis into early endosomes that is dynamin dependent.[58]

In an elegant follow-up report, the synergistic aspects of engaging both CD22 and BCR simultaneously and the resulting perturbation on BCR downstream signaling were studied with glycoprotein mimics.[59] Taking advantage of a well-defined, ROMP-derived N-hydroxy succinimidyl (NHS) ester polymer (DOP = 250), the backbone was elaborated with either BCR ligands (DNP), CD22 ligands (CD22L) or both (DNP/CD22L). By tethering both BCR and CD22 ligands to the same polymer backbone, the dynamics of BCR and CD22 would be restricted to membrane regions enriched with the two transmembrane proteins. Consistent with previous observations, CD22L glycopolymers did not bind to B cell surfaces, which is further evidence that the native CD22L glycan structure does not displace endogenous *cis* CD22 interactions. As anticipated, both DNP-containing polymers adhered to B cell surfaces. To examine the regulatory consequences of clustering both BCR and CD22, calcium ion influx was monitored by a ratiometric calcium-chelating dye as a proxy for BCR signaling. Cells treated with DNP polymer displayed an increase in cytosolic calcium ion concentration. The DNP/CD22L copolymer, on the other hand, had no such effect. No calcium influx was observed for the copolymer, and downstream BCR signaling targets, such as Syk and PLC γ2, had reduced phosphorylation compared with DNP polymer-treated cells (Figure 1.3). Interestingly, incubation with the DNP polymer *or* the DNP/CD22L copolymer produced similar early signaling patterns in B cells; both Lyn and CD22 became phosphorylated upon polymer engagement. The authors proposed that after BCR attachment, the increase in local concentration of polymer on the membrane facilitates CD22 *trans* association. Collectively, the data also indicate that co-clustering of BCR and CD22 on the cell surface elicits early B cell activation but attenuates downstream signaling effectors. The copolymer data strongly suggest that CD22 dynamics can be modulated by the glycosylated state of antigen, which could be an innate form of self-recognition. This insight would have been difficult to obtain without the aid of homogeneous glycoprotein mimics.

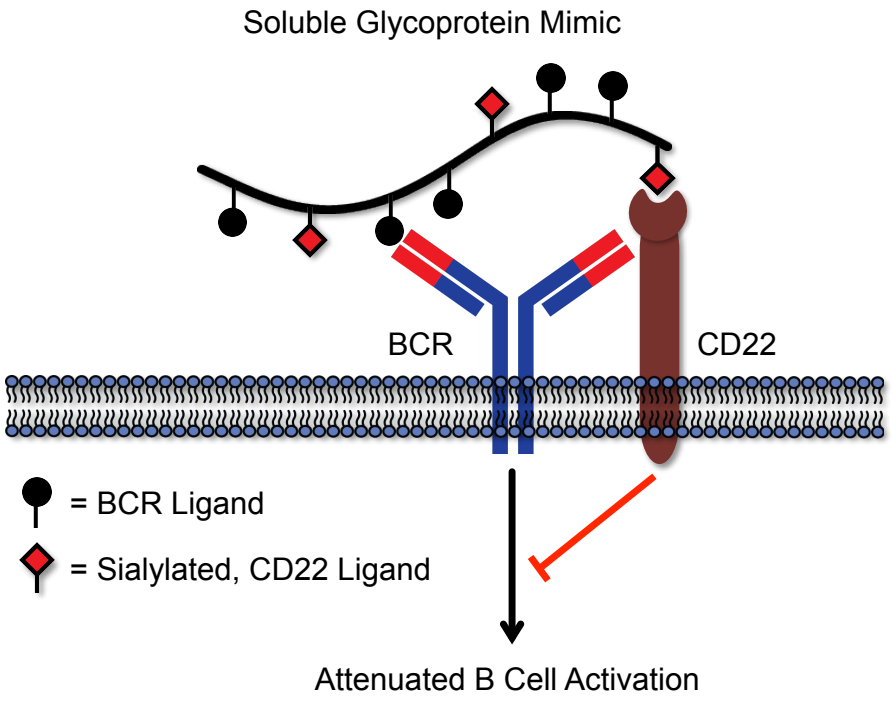


Figure 1.3. A soluble copolymer, glycoprotein mimic engages both BCR and CD22 simultaneously. Glycopolymer binding reorganizes CD22 to sites of BCR activation and blocks downstream signaling. The CD22 ligands are composed of sialic acid residues, suggesting that sialylation may be a molecular form of 'self' in mammals.

To further explore the influence of B cell siglecs, such as CD22 and Siglec-G, on tolerance in the context of a living organism's immune system, polyacrylamide (PA) copolymers were synthesized bearing both a BCR ligand (NP) and either a native CD22 ligand (NeuGc) or a high-affinity, unnatural CD22 ligand (bNeuGc).[60] Mice immunized with NP-PA produced a strong antibody, IgM and IgG3, response as expected. In contrast, mice immunized with either the NP-PA-NeuGc copolymer or the NP-PA-bNeuGc copolymer failed to elicit high anti-NP titers. These results agree with the previous cell-based findings. To test whether glycopolymer antigen challenge shaped the organism's subsequent immune response, mice that had been treated with sialylated antigen were then immunized with unsialylated NP-PA. Drastically reduced antibody titer was observed, a hallmark of antibody suppression and immune tolerance. Tolerance induction was associated with suboptimal B cell proliferation and arrest of plasma cell differentiation. Additionally, long-term tolerance was correlated with the strength of siglec binding as NP-PA-bNeuGc copolymers produced the highest immune tolerance. CD22 and Siglec-G can thus confer a graded output when presented with different glycosylated antigens.

Yet, what mechanisms are responsible for producing a spectrum of B cell activation? Kiessling and coworkers decided to explore how the dynamics of lectin trafficking could be responsible for B cell signaling regulation.[61] To accomplish this, B cells were incubated with DNP polymer or DNP/CD22L copolymers, and BCR endocytosis was tracked by fluorescence microscopy. The cell surface half-life of BCR was significantly longer for cells treated with DNP polymer compared to cells treated with DNP/CD22L copolymer. Co-clustering of BCR and CD22 on B cell surfaces triggered endocytosis, which resulted in accumulation of BCR in early

endosomes after DNP/CD22L copolymer incubation. Glycosylation-dependent internalization was mediated by dynamin and initiated by early phosphorylation of BCR and CD22 through Lyn. This mechanism of BCR regulation provides a means for B cell's to tailor their BCR levels based on glycoprotein antigen and in turn increase the threshold of B cell activation to subsequent antigen encounter.

Through the use of soluble, sialylated glycoprotein mimics, the intimate details of B cell siglecs have emerged. Glycopolymers have served as homogeneously glycosylated replacements for glycoprotein antigens in these experiments. As a result, a picture of sialylation as a form of "self" has been partly resolved mechanistically; copolymers presenting both BCR and CD22 ligands also provided evidence that self-recognition can lead to long-term immune tolerance *in vivo*.

Immune cells other than B cells also possess unique repertoires of siglecs. One example is natural killer (NK) cells that act as early responders to foreign entities, leading to cell death and clearance.[62] NK cells prominently express Siglec-7 and to a lesser extant Siglec-9.[63] Early work by Nicoll *et al.* used soluble glycopolymers to establish that similar to CD22, Siglec-7 is indeed masked on NK cells by the disialylated ganglioside GD3.[64] But how siglecs on NK cells respond during cell-cell contact had not been fully elucidated. To engineer *cell surfaces* with chemically defined glycoprotein mimics, the Bertozzi lab developed a strategy to synthesize a glycopolymer functionalized with a lipid at one terminus.[65] This enabled the passive insertion of glycopolymers into cell membranes and expanded the potential to not only investigate secreted glycoproteins but membrane-bound glycoproteins as well. Hudak *et al.* harnessed this technology to modify targets cells of NK cell-mediated lysis.[66] As an integral part of the innate immune response, NK cells act to contain the growth and metastasis of tumor cells. Aggressive cancers though are able to evade NK cells through the release of activating receptors or expression of inhibitory motifs. One such motif is hypersialylation. The prevailing assumption was that cell surface sialylation on cancer cells presents a physical barrier to NK cell engagement. But owing to the presence of siglecs on NK cell surfaces, lectin binding had also been hypothesized to contribute to NK cell protection. To isolate the molecular basis of NK cell inhibition, a lipid-functionalized poly(methylvinylketone) polymer was elaborated with a large panel of aminooxy glycan structures. Jurkat cells, a hyposialylated cell line that are known to be susceptible to NK cell killing, were decorated with polymers and then incubated with either purified NK cells or peripheral blood mononuclear cells (PBMC). NK cytotoxicity of Jurkat cells was strongly inhibited by sialic acid containing glycopolymers but not polymers lacking the terminal sugar or polymers possessing similar charge, acetic acid, or polarity, glycerol. Function-blocking antibodies were then used to prevent any specific interactions occurring with cell surface sialic acid residues. The anti-Siglec-7, but not the anti-Siglec-9, antibody completely abolished any cytotoxicity protection that sialic acid polymers afforded. With this striking finding in hand, fluorescence micrographs were taken and revealed that Jurkat cells displaying the sialic acid glycopolymer caused the reorganization of Siglec-7 on the surface of NK cells. Initially, Siglec-7 is distributed uniformly on the surface, but after engagement the lectin is concentrated at cell-cell synapses. Subsequent immunoprecipitation of NK Siglec-7 showed glycopolymer-dependent phosphorylation and SHP recruitment, which points to phosphatases as the signaling regulator of NK cell degranulation. Based on results from glycoengineered cells, the authors concluded that hypersialylation of aggressive cancer cells is a form of microevolutionary immunoevasion that ultimately changes Siglec-7 dynamics on NK cells.

Galectins, another family of carbohydrate binding proteins, differ considerably in topology from selectins and siglecs.[67] These proteins are soluble, and several family members are thought to have divergent functions both extracellularly and intracellularly. Whereas siglecs and selectins bind sialylated structures, galectins recognize a different class of glycans, *N*-acetyllactosamine residues. Moreover, the prototype and chimera-type galectins have the ability to self-assemble in solution while the tandem repeat-type galectins have two carbohydrate recognition domains (CRDs) that are fused together through a flexible linker peptide. Stemming from their oligomeric nature, it was proposed that galectins may form a lattice-like structure with glycoproteins on cell surfaces.[68] These galectin microdomains could act as adaptors, assembling proteins and lipids based on their glycosylation state, and modulate the biochemical activity of glycoprotein ligands. Work by Dennis and coworkers provided experimental support for this model with Mgat5-knockdown cells and mice lacking the N-acetylglucosaminyltransferase responsible for adding LacNAc branches onto N-glycans.[69] Limiting the number of galectin ligands on cell surface glycoproteins altered the internalization rates of a number of growth factor receptors, thereby indicating that galectin lattices may extend the cell surface residence time of certain glycoproteins. The observed increase in surface half-life of the glycoprotein receptors in turn affected the duration and thus the strength of signal transduction in response to growth factor binding.

Unfortunately, direct experimental validation of galectin-glycoprotein lattices had eluded researchers. Without a means of detection, physical properties, such as size and lifetime, were largely undefined for galectin-mediated aggregation, which impeded further studies on the differences between family members in maintaining membrane organization. To address this experimental deficiency, membrane-anchored glycopolymers and fluorescence spectroscopy were leveraged to verify the existence of galectin lattices.[70] One requirement of fluorescence spectroscopy, a time-dependent technique well suited for studying dynamic systems, is the well-defined placement and stoichiometry of a fluorescent dye within a macromolecular architecture. Reversible addition-fragmentation chain transfer (RAFT) polymerization, another type type of living polymerization, offers the opportunity to individually manipulate both ends of a polymer chain while also affording predictable DOP.[71,72] Lactosyl and cellobiosyl (negative control) acrylamide monomers were therefore subjected to RAFT polymerization conditions in the presence of a lipid-functionalized chain transfer agent. Fluorescent glycoprotein mimics were prepared by end group modification, yielding glycopolymers containing either a single FRET donor or acceptor fluorophore. After embedding donor and acceptor glycopolymers (1:1) into living *ldlD* CHO cell membranes, fluorescence lifetime of the donor fluorophore was monitored in real-time for a period of 40 min. A decrease in fluorescence lifetime of the donor-functionalized lactosyl polymer was observed in the presence of galectin-1, but inhibiting galectin-1 binding or changing the polymer's glycans (cellobiosyl polymers) abolished the FRET signature. Moreover, the diffusion time, a parameter that can be measured by Fluorescence Correlation Spectroscopy and related to clustering, for the lactosyl glycopolymer increased only in the presence of galectin-1. By examining the data in real time, it appears that galectin-1 mediates a dynamic cross-linking of membrane glycoprotein mimics on the minutes timescale and results in extended galectin networks, which are themselves stable for tens of minutes.

Now considered invaluable tools in the chemical biologist's toolkit, glycoprotein mimics have been utilized *in vivo* to understand the flow of information from a glycosylated biomolecule to a lectin and the resulting cellular decision-making that is altered by carbohydrate binding. What has emerged is a picture of lectins whose functions are highly dependent on changes in

dynamics that occur upon ligand binding. Following glycoprotein mimic association, lectins redistribute, form signaling clusters, and process signals, all of which are regulated by the nature of the glycan partners themselves. Glycoprotein mimics have been instrumental in defining the parameters that affect lectin-mediated signaling and membrane organization both in cell culture and in organismal studies. Present mimics, however, lack the signaling domains associated with many glycoprotein ligands. In the future, it would be interesting to combine mucin-type polymeric structures with functional domains through semi-synthesis. Using this approach, questions regarding the feedback between lectin dynamics and glycoprotein function can start to be assessed.

Site-selective Tagging and Imaging Technologies

Proteins are often organized in biological systems on both the molecular and supramolecular scale. Changes in conformation or oligomeric state have significant implications for a protein's role in carrying out a specified set of functions in the cell. These different configurations can completely alter the binding properties of proteins, acting like on-off switches in some cases. An interesting example is actin, which was once thought to have a simple well-defined function and organization. Now there is ample evidence that it can exist as a soluble monomer in the cytoplasm/nucleus or assemble dynamically to form different oligomeric and polymeric species, such as fibers and bundles, all of which have vastly different cellular effects: monomeric nuclear actin can associate with transcription regulators and control gene expression, whereas filamentous actin acts as scaffolding for trafficking and cell motility.[73] As demonstrated in the sections above, lectin organization and dynamics have a surprising, nontrivial influence on cellular function. Yet, imaging the molecular details of lectin organization in a complex biological sample and following lectin dynamics are major challenges. First, to visualize a single protein molecule with optics requires a means to "break" the diffraction limit of light since the length scale of lectins is typically 5-10 nm, well below the diffraction limit, ~200 nm. Protein dynamics further complicates observation due to the short timescales associated with molecular diffusion and polypeptide conformational changes. These features necessitate fast detectors, i.e. hardware, to capture such events. Second, most proteins do not possess tags that distinguish their location from others in the dense, crowded cytoplasm or extracellular matrix. Labeling with fluorescent protein fusions or contrast agents can aid in visualization, but CRDs are small and globular and selecting the proper technique is of the utmost importance to avoid perturbing the lectin's intrinsic characteristics. Advances in both protein modification and imaging methodology and hardware have recently begun to address the visualization of lectin organization. In the following section, we outline the very few examples in the literature that focus on mammalian lectins. Our intent is to spark additional research in this burgeoning area.

Leukocytes home to sites of damage and inflammation through L-selectin bonds with sialylated glycans on endothelial cells. This phenomenon is surprising because the attachment of leukocytes is enhanced under flow. While certainly a counterintuitive finding, it was originally rationalized by changes in cell surface area or number of adhesions under shear force. But to rigorously examine the molecular basis of leukocyte trafficking, a technique capable of observing short adhesion events was required. Yago *et al.* met this challenge by immobilizing the L-selectin ligand, P-selectin glycoprotein ligand-1 (PSGL-1), on a substrate and measuring the biophysical parameters that govern adhesion under flow with fast microscopy detectors (250-500 frames per second).[74] The authors monitored off-rates (k_{off}), among other parameters, for the L-

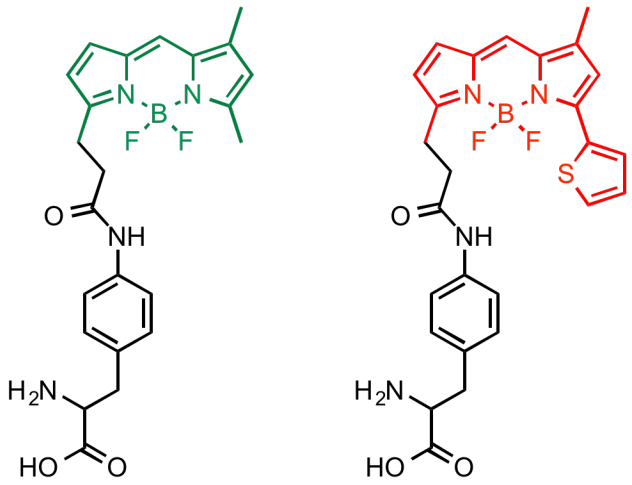
selectin-PSGL-1 interaction in neutrophils under various tether forces. They observed decreases in k_{off} as shear force was increased. This behavior is known as a catch bond, where bond lifetime lengthens under force. In this case, advances in instrumentation were pivotal for assigning the first biological function of catch bonds, i.e. flow enhancement of leukocyte attachment. This work also provides credence to the notion that lectin dynamics, e.g. force-induced conformation changes, are important for large-scale biological transformations, like leukocyte rolling and extravasation.

DC-SIGN, as previously discussed, is a C-type lectin that is involved in antigen uptake and processing in dendritic cells. Its cell surface organization has garnered significant attention due to the fact that DC-SIGN is involved in recognizing glycan epitopes on the surface of large pathogens, including Human Immunodeficiency Virus (HIV), Ebola virus, *Candida albicans*, *Leishmania* and others. In fact, DC-SIGN microdomains were first detected using wide-field microscopy and estimated to be between 300 and 1500 nm in size. Given the optical resolution of traditional microscopy and dimensions of viral particles (~80-100 nm), researchers were interested in whether DC-SIGN formed smaller clusters on the surface of immature dendritic cells. Super-resolution microscopy is a suite of techniques that allows spatial localization of individual molecules with 10's of nm resolution. To examine DC-SIGN's organization, Blink Microscopy, a super-resolution method, was creatively applied to estimate the cluster size of DC-SIGN.[75] Blink Microscopy relies on driving the majority of fluorophores in a densely-labeled sample into a dark state through a combination of redox reagents, leaving a few dyes fluorescing for short *on* times (~ms). By fitting the point spread function, the centroid of the fluorophore can be calculated and precisely localized with subdiffraction resolution. DC-SIGN was imaged using immunofluorescence with an ATTO655 dye that can be cycled in and out of a dark state in the presence of ascorbic acid and *N,N*-methylviologen. DC-SIGN appeared to be organized into nanodomains with a mean full width at high maximum (FWHM) value of 76 nm. By analyzing the number of blink localizations per cluster, the occupancy was calculated to be ~12 molecules, or three tetramers of DC-SIGN, indicating that the cluster was sparsely occupied. This finding was also consistent with the ability of lipid molecules to easily diffuse through DC-SIGN membrane domains. The multivalency associated with DC-SIGN and its limited density in clusters suggests an effective mechanism to ensure entry of weakly binding antigens that only require a single interaction for uptake, e.g. HIV.

Galectin-3 (Gal-3) is the only member of the galectin family that belongs to the chimera subtype. This protein has been reported to oligomerize *in vitro* upon binding glycan ligands, forming pentamers in solution.[13] The N-terminal stalk region of Gal-3 is thought to be responsible for and the site of self-assembly between individual Gal-3 monomers. This presents a unique problem, though, for tracking Gal-3's oligomerization in a biological sample. If both the N-terminus and the C-terminus, which contains the carbohydrate recognition domain, are absolutely essential for Gal-3's function, then tagging Gal-3 with a large fluorescent protein fusion would disrupt critical interactions at either of these domains. Despite this limitation, Nieminen *et al.* were still interested in observing Gal-3 oligomerization using FRET in the context of neutrophil-endothelial cell interfaces.[76] While FRET is a suitable reporter of Gal-3 assembly, the question of Gal-3 modification remained. To address this, a *small* organic donor or acceptor fluorophore, Alexa Fluor 488- or Alexa Fluor 555-cadaverine, was conjugated to the C-terminus of Gal-3 in a site-specific manner by the enzyme, transglutaminase. The peptide tag for transglutaminase, 7 amino acids in this example, did not alter Gal-3's glycan binding properties even after fluorophore ligation. By precisely placing the dye at end of Gal-3's primary structure,

robust FRET signals were observed in unprimed neutrophils and in the context of neutrophil-endothelial cell adhesions. On the other hand, Gal-3 monomers modified non-specifically at randomly distributed lysine residues showed limited FRET. Together, these results suggest that Gal-3 is oligomerized within the glycocalyx on cell surfaces. This study also emphasizes the unmet need and demand for developing methods capable of site-selectively modifying lectins without perturbing activity or quaternary lectin structure. For instance, in very recent work, Gal-3 has been tethered with a SNAP-tag that allows for modification with a myriad of benzylguanine derivatives, opening the door for a range of markers useful for observing lectin dynamics/organization.[77]

Finally, fluorescent amino acids are beneficial from an imaging perspective since they are small in size and can be placed at many points *throughout* the protein backbone - characteristics that are important for following lectin dynamics. Fluorescence is also extremely sensitive to the local environment and has been used extensively to detect binding events *in vivo*. Based on these advantages, Hohsaka and coworkers recently pioneered the incorporation of fluorescent amino acids into lectins.[78,79] Through both cell-free and amber stop codon suppression methods, BODIPY fluorescent amino acids were efficiently installed in maltose-binding protein as well as in a sialic acid-binding protein (Figure 1.4). In both examples, placement of the fluorophore near the glycan-binding site created a fluorogenic lectin, i.e. fluorescence of the amino acid is quenched by a nearby tryptophan residue in the absence of ligand, but when maltose or sialyllactose was administered fluorescence turns "on." These examples are encouraging since they suggest that fluorescent amino acids may be well suited to visualize small lectins *in vivo*, obfuscating addition of protein fusions or peptide tags to a lectin's sequence. They also illustrate a path to selectively visualize lectins that are engaged with their glycoprotein ligands as opposed to 'free' protein.



BODIPY-FL-aminophenylalanine **BODIPY-558-aminophenylalanine**

Figure 1.4. Fluorescent amino acids that have been incorporated into maltose-binding protein and a sialic acid-binding protein.

Chemical tools have revolutionized the way molecules are imaged and tracked inside complex biological milieu. Many of these techniques are just now being applied to lectins.

Lectins assemble, disassemble, form cross-linked 'bridges' between glycoproteins and adjacent cells, and generally act in complex, multimeric forms. Non-traditional methods are needed to observe the debated details of these processes *in vivo*. Recently, new hardware and chemical reagents have revealed, for instance, the nature of lectin binding under force and the spatial organization of lectin domains in membranes. We believe that by leveraging new contributions in protein modification and better imaging hardware/techniques our present understanding of lectin dynamics and binding behavior will give way to a more refined description. As well, we hope that in the future researchers will take advantage of single molecule techniques, such as single particle tracking, coupled with site-specific fluorophore-lectin conjugates to resolve lectin dynamics in different subcellular compartments, an unexplored frontier.

Conclusion

Chemists have transformed the field of glycobiology by contributing a number of methods to interrogate mammalian lectin dynamics *in vivo*. These tools have come in a variety of molecular forms, ranging from small molecule metabolites to large macromolecular agents. Cross-linking sugars have revealed extraordinary features of CD22 that were wholly inaccessible using typical molecular biology techniques. The fact that CD22 is segregated into homomultimers in resting B cells has shed light on the *in vivo* consequences of lectin reorganization for tuning immune cell activation. Glycoprotein mimics have, for instance, clarified a novel glycan-based mechanism of NK cytotoxicity evasion by cancer cells. Escaping immunosurveillance was, not surprisingly, predicated on glycan-siglec interactions. And new imaging methods, such as super-resolution microscopy, have settled long-standing questions concerning DC-SIGN membrane domains. By precisely locating individual DC-SIGN molecules, DC-SIGN clusters were found to be composed of only 12 molecules, which explains their ability to efficiently mobilize and capture viral particles. However, *in vivo* specificity is still undefined for the vast majority of lectins and only recently have researchers begun to appreciate that supramolecular assembly and subcellular dynamics of lectins profoundly dictate function and cell behavior. We believe that chemical tools will be indispensible for elucidating how lectins interpret glycomic changes in the future. Our hope is that this review will stimulate further interest and work in this new, exciting area at the interface of chemistry and biology.

References

(1) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. *Essentials of Glycobiology*; 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor N.Y., 2008.

(2) Ohtsubo, K.; Marth, J. D. *Cell* **2006**, *126*, 855–867.

(3) Sharon, N. *J. Biol. Chem.* **2007**, *282*, 2753–2764.

(4) Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *28*, 321–327.

(5) Sharon, N.; Lis, H. *Glycobiology* **2004**, *14*, 53R–62R.

(6) Paulson, J. C.; Blixt, O.; Collins, B. E. *Nat. Chem. Biol*. **2006**, *2*, 238–248.

(7) Price, J. L.; Culyba, E. K.; Chen, W.; Murray, A. N.; Hanson, S. R.; Wong, C.-H.; Powers, E. T.; Kelly, J. W. *Biopolymers* **2012**, *98*, 195–211.

(8) Wormald, M. R.; Dwek, R. A. *Structure* **1999**, *7*, R155–160.

(9) Collins, B. E.; Paulson, J. C. *Curr. Opin. Chem. Biol.* **2004**, *8*, 617–625.

(10) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555–578.

(11) Colin Hughes, R. *Curr. Opin. Struct. Biol.* **1992**, *2*, 687–692.

(12) Brewer, C. F.; Miceli, M. C.; Baum, L. G. *Curr. Opin. Struct. Biol.* **2002**, *12*, 616–623.

(13) Ahmad, N.; Gabius, H.-J.; André, S.; Kaltner, H.; Sabesan, S.; Roy, R.; Liu, B.; Macaluso, F.; Brewer, C. F. *J. Biol. Chem.* **2004**, *279*, 10841 –10847.

(14) Bourne, Y.; Bolgiano, B.; Liao, D.-I.; Strecker, G.; Cantau, P.; Herzberg, O.; Feizi, T.; Cambillau, C. *Nat. Struct. Biol.* **1994**, *1*, 863–870.

(15) Kremers, G.-J.; Gilbert, S. G.; Cranfill, P. J.; Davidson, M. W.; Piston, D. W. *J. Cell. Sci.* **2011**, *124*, 157–160.

(16) Kayser, H.; Zeitler, R.; Kannicht, C.; Grunow, D.; Nuck, R.; Reutter, W. *J. Biol. Chem.* **1992**, *267*, 16934–16938.

(17) Kayser, H.; Geilen, C. C.; Paul, C.; Zeitler, R.; Reutter, W. *FEBS Lett.* **1992**, *301*, 137–140.

(18) Keppler, O. T.; Horstkorte, R.; Pawlita, M.; Schmidt, C.; Reutter, W. *Glycobiology* **2001**, *11*, 11R–18R.

(19) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125–1128.

(20) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007–2010.

(21) Hang, H. C.; Yu, C.; Kato, D. L.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14846–14851.

(22) Vocadlo, D. J.; Hang, H. C.; Kim, E.-J.; Hanover, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9116–9121.

(23) Rabuka, D.; Hubbard, S. C.; Laughlin, S. T.; Argade, S. P.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2006**, *128*, 12078–12079.

(24) Beahm, B. J.; Dehnert, K. W.; Derr, N. L.; Kuhn, J.; Eberhart, J. K.; Spillmann, D.; Amacher, S. L.; Bertozzi, C. R. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 3347–3352.

(25) Belardi, B.; de la Zerda, A.; Spiciarich, D. R.; Maund, S. L.; Peehl, D. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 14045–14049.

(26) Haga, Y.; Ishii, K.; Hibino, K.; Sako, Y.; Ito, Y.; Taniguchi, N.; Suzuki, T. *Nat. Commun.* **2012**, *3*, 907.

(27) Lin, W.; Du, Y.; Zhu, Y.; Chen, X. *J. Am. Chem. Soc.* **2014**, *136*, 679–687.

(28) Pham, N. D.; Parker, R. B.; Kohler, J. J. *Curr. Opin. Chem. Biol.* **2013**, *17*, 90–101.

(29) Walker, J. A.; Smith, K. G. C. *Immunology* **2008**, *123*, 314–325.

(30) Tedder, T. F.; Poe, J. C.; Haas, K. M. *Adv. Immunol.* **2005**, *88*, 1–50.

(31) Han, S.; Collins, B. E.; Bengtson, P.; Paulson, J. C. *Nat. Chem. Biol.* **2005**, *1*, 93–97.

(32) Ramya, T. N. C.; Weerapana, E.; Liao, L.; Zeng, Y.; Tateno, H.; Liao, L.; Yates, J. R.; Cravatt, B. F.; Paulson, J. C. *Mol. Cell. Proteomics* **2010**, *9*, 1339–1351.

(33) Yu, H.; Chen, X. *Org. Biomol. Chem.* **2007**, *5*, 865–872.

(34) Tanaka, Y.; Kohler, J. J. *J. Am. Chem. Soc.* **2008**, *130*, 3278–3279.

(35) Bond, M. R.; Zhang, H.; Kim, J.; Yu, S.-H.; Yang, F.; Patrie, S. M.; Kohler, J. J. *Bioconjug. Chem.* **2011**, *22*, 1811–1823.

(36) Yu, S.-H.; Boyce, M.; Wands, A. M.; Bond, M. R.; Bertozzi, C. R.; Kohler, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4834–4839.

(37) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. P *Nat. Methods* **2005**, *2*, 201–206.

(38) Poupon, E.; Nay, B. *Biomimetic Organic Synthesis*; John Wiley & Sons, 2011.

(39) Ladmiral, V.; Melia, E.; Haddleton, D. M. *Eur. Polym. J.* **2004**, *40*, 431–449.

(40) Turnbull, W. B.; Stoddart, J. F. *J. Biotechnol.* **2002**, *90*, 231–255.

(41) De la Fuente, J. M.; Penadés, S. *Biochim. Biophys. Acta* **2006**, *1760*, 636–651.

(42) Schnaar, R. L.; Lee, Y. C. *Biochemistry* **1975**, *14*, 1535–1541.

(43) Sigal, G. B.; Mammen, M.; Dahmann, G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996**, *118*, 3789–3800.

(44) Spain, S. G.; Gibson, M. I.; Cameron, N. R. *J. Polym. Sci. A Polym. Chem.* **2007**, *45*, 2059–2072.

(45) Becer, C. R. *Macromol. Rapid Commun.* **2012**, *33*, 742–752.

(46) Deniaud, D.; Julienne, K.; Gouin, S. G. *Org. Biomol. Chem.* **2011**, *9*, 966–979.

(47) Kiessling, L. L.; Grim, J. C. *Chem. Soc. Rev.* **2013**, *42*, 4476–4491.

(48) Rosen, S. D. *Ann. Rev. Immunol.* **2004**, *22*, 129–156.

(49) Mowery, P.; Yang, Z.-Q.; Gordon, E. J.; Dwir, O.; Spencer, A. G.; Alon, R.; Kiessling, L. L. *Chem. Biol.* **2004**, *11*, 725–732.

(50) Matyjaszewski, K. Fundamentals of Controlled/Living Radical Polymerization. In *Encyclopedia of Radicals in Chemistry, Biology and Materials*; John Wiley & Sons, Ltd, 2012.

(51) Sutthasupa, S.; Shiotsuki, M.; Sanda, F. *Polym J* **2010**, *42*, 905–915.

(52) Liu, S.; Kiick, K. *Polym. Chem.* **2011**, *2*, 1513–1522.

(53) Švajger, U.; Anderluh, M.; Jeras, M.; Obermajer, N. *Cell. Signal.* **2010**, *22*, 1397–1405.

(54) Prost, L. R.; Grim, J. C.; Tonelli, M.; Kiessling, L. L. *ACS Chem. Biol.* **2012**, *7*, 1603–1608.

(55) Denda-Nagai, K.; Aida, S.; Saba, K.; Suzuki, K.; Moriyama, S.; Oo-Puthinan, S.; Tsuiji, M.; Morikawa, A.; Kumamoto, Y.; Sugiura, D.; Kudo, A.; Akimoto, Y.; Kawakami, H.; Bovin, N. V.; Irimura, T. *J. Biol. Chem.* **2010**, *285*, 19193–19204.

(56) Collins, B. E.; Blixt, O.; Han, S.; Duong, B.; Li, H.; Nathan, J. K.; Bovin, N.; Paulson, J. C. *J. Immunol.* **2006**, *177*, 2994–3003.

(57) Zaccai, N. R.; Maenaka, K.; Maenaka, T.; Crocker, P. R.; Brossmer, R.; Kelm, S.; Jones, E. Y. *Structure* **2003**, *11*, 557–567.

(58) Tateno, H.; Li, H.; Schur, M. J.; Bovin, N.; Crocker, P. R.; Wakarchuk, W. W.; Paulson, J. C. *Mol. Cell. Biol.* **2007**, *27*, 5699–5710.

(59) Courtney, A. H.; Puffer, E. B.; Pontrello, J. K.; Yang, Z. Q.; Kiessling, L. L. *Proc. Natl. Acad. Sci.* **2009**, *106*, 2500.

(60) Duong, B. H.; Tian, H.; Ota, T.; Completo, G.; Han, S.; Vela, J. L.; Ota, M.; Kubitz, M.; Bovin, N.; Paulson, J. C.; Paulson, J.; Nemazee, D. *J. Exp. Med.* **2010**, *207*, 173–187.

(61) Courtney, A. H.; Bennett, N. R.; Zwick, D. B.; Hudon, J.; Kiessling, L. L. *ACS Chem. Biol.* **2014**, *9*, 202–210.

(62) Kumar, V.; McNerney, M. E. *Nat. Rev. Immunol.* **2005**, *5*, 363–374.

(63) Crocker, P. R.; Paulson, J. C.; Varki, A. *Nat. Rev. Immunol.* **2007**, *7*, 255–266.

(64) Nicoll, G.; Avril, T.; Lock, K.; Furukawa, K.; Bovin, N.; Crocker, P. R. *Eur. J. Immunol.* **2003**, *33*, 1642–1648.

(65) Rabuka, D.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 5947–5953.

(66) Hudak, J. E.; Canham, S. M.; Bertozzi, C. R. *Nat Chem Biol* **2014**, *10*, 69–75.

(67) Di Lella, S.; Sundblad, V.; Cerliani, J. P.; Guardia, C. M. A.; Estrin, D. A.; Vasta, G. R.; Rabinovich, G. A. *Biochemistry* **2011**.

(68) Rabinovich, G. A.; Toscano, M. A.; Jackson, S. S.; Vasta, G. R. *Curr. Opin. Struct. Biol.* **2007**, *17*, 513–520.

(69) Partridge, E. A.; Le Roy, C.; Di Guglielmo, G. M.; Pawling, J.; Cheung, P.; Granovsky, M.; Nabi, I. R.; Wrana, J. L.; Dennis, J. W. *Science* **2004**, *306*, 120 –124.

(70) Belardi, B.; O'Donoghue, G. P.; Smith, A. W.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9549–9552.

(71) Lowe, A. B.; McCormick, C. L. *Prog. Polym. Sci.* **2007**, *32*, 283–351.

(72) Keddie, D. J. *Chem. Soc. Rev.* **2014**, *43*, 496.

(73) Olson, E. N.; Nordheim, A. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 353–365.

(74) Yago, T.; Wu, J.; Wey, C. D.; Klopocki, A. G.; Zhu, C.; McEver, R. P. *J. Cell Biol.* **2004**, *166*, 913–923.

(75) Itano, M. S.; Steinhauer, C.; Schmied, J. J.; Forthmann, C.; Liu, P.; Neumann, A. K.; Thompson, N. L.; Tinnefeld, P.; Jacobson, K. S. *Biophys. J.* **2012**, *102*, 1534–1542.

(76) Nieminen, J.; Kuno, A.; Hirabayashi, J.; Sato, S. *J. Biol. Chem.* **2006**, *282*, 1374–1383.

(77) Kupper, C. E.; Böcker, S.; Liu, H.; Adamzyk, C.; van de Kamp, J.; Recker, T.; Lethaus, B.; Jahnen-Dechent, W.; Neuss, S.; Müller-Newen, G.; Elling, L. *Curr. Pharm. Des.* **2013**, *19*, 5457–5467.

(78) Iijima, I.; Hohsaka, T. *Chembiochem* **2009**, *10*, 999–1006.

(79) Ito, Y.; Hohsaka, T. *Bull. Chem. Soc. Jpn.* **2013**, *86*, 729–735.

Chapter 2

Well-Defined Glycoprotein Mimics for Exploring Galectin Recognition

Portions of this chapter were modified from: Belardi, B.; O'Donoghue, G. P.; Smith, A. W.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9549 - 9552.

Chapter 2. Well-Defined Glycoprotein Mimics for Exploring Galectin Recognition

Introduction

Elucidating the diverse functions of the glycome, the collection of all complex carbohydrate structures expressed in cells and tissues, is one of the most difficult and sought-after challenges that has emerged in the post-genomic era.[1-3] Intricate glycan structures adorn the exterior of the cell membrane and mediate cellular interactions with circulating proteins, extracellular matrix components, pathogens, and other cells. These interactions have physiological consequences in terms of diseases, inflammatory responses, and differentiation of various cell types.[4-14] The ability to modulate such extracellular interactions would facilitate our understanding of the roles cell surface glycans play in various biological processes.

Galectins are one family of glycan-binding proteins that recognize N- and O-linked glycans bearing the terminal disaccharide, *N*-acetyllactosamine.[15] Characterizing the organizational, signaling, and cell fate consequences of galectin binding is a difficult task in the context of a complex biological sample. Galectin ligands form a set of secreted and membrane-bound glycoproteins that have numerous cellular functions, many of which are critical to cellular homeostasis.[16] These galectin-binding glycoproteins are characterized by their heterogeneity: a glycoprotein exists in many different glycoforms that differ from protein-to-protein as well as from site-to-site on a single peptide backbone. To study specific glycan-galectin interactions *in vivo* requires a technique to tailor the glycan structures appended to glycoproteins. Unfortunately, glycosylation, a post-translational modification, is refractory to genetic manipulations, and perturbing the cellular glycomic profile globally can affect other properties associated with proteins, e.g. folding and stability, that are not related to galectin binding.[17,18]

Synthetic glycopolymers have proven to be powerful functional surrogates for natural glycoconjugates, particularly in situations where the complexity and heterogeneity of the native biomolecules undermines experimental inquiry.[19-24] For decades, chemists have made use of various glycopolymer architectures to study glycan-receptor interactions related to the immune response,[25,26] viral infection,[27,28] bacterial signaling[29] and neurobiology,[30] and recently we employed synthetic glycopolymers as ligands for microarray-based studies of glycan binding[31,32] proteins and for cell-surface functionalization[33].

We sought to emulate the features of native glycoproteins by constructing glycopolymer-based mimics capable of binding galectins. Membrane glycoproteins possess multiple domains that are linked to one another in a linear fashion. MUC1, an O-linked mucin glycoprotein, for instance, consists of a a cytosolic recognition domain, transmembrane domain, an SEA domain, and a heavily glycosylated, repeating motif.[34] Glycopolymers were designed to mirror this feature and possess the following attributes (Figure 2.1): (1) a lipid anchor at one terminus for membrane insertion and display of the polymers on cell surfaces, (2) a polymer backbone containing pendant glycans that are specific for galectin, and (3) a tag at the other end for tracking mimics in the biological milieu. Herein, we describe the development of a living polymerization method enabling the facile dual end-functionalization of a well-defined glycopolymer.

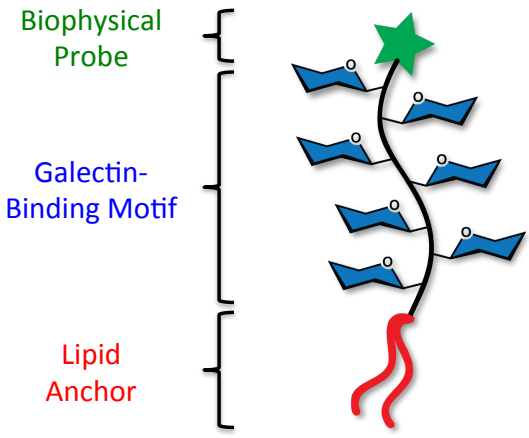


Figure 2.1. Glycoprotein mimic design. A well-defined glycopolymer differentially functionalized at both termini is targeted. The mimic possesses the ability to insert into cell membranes and bind galectins.

Results and Discussion

Conventional free radical polymerization (FRP) suffers from recombination and chain transfer processes, resulting in polymers with uncontrollable lengths and high polydispersities. Using FRP, there is no obvious means to form linear polymers with two different end groups. However, a number of living polymerization methods involving radical propagation, known as controlled radical polymerizations (CRP), have recently been developed that minimize termination events during the polymerization reaction.[35] This approach affords tunable chain lengths and uniformly distributed polymers as well as the opportunity to tailor the polymer with different functional groups at the termini. We thus envisioned controlling the polymerization of a glycosyl monomer with a lipid-functionalized species capable of producing a multi-domain linear polymer.

Synthesis of a lactosyl methacrylate

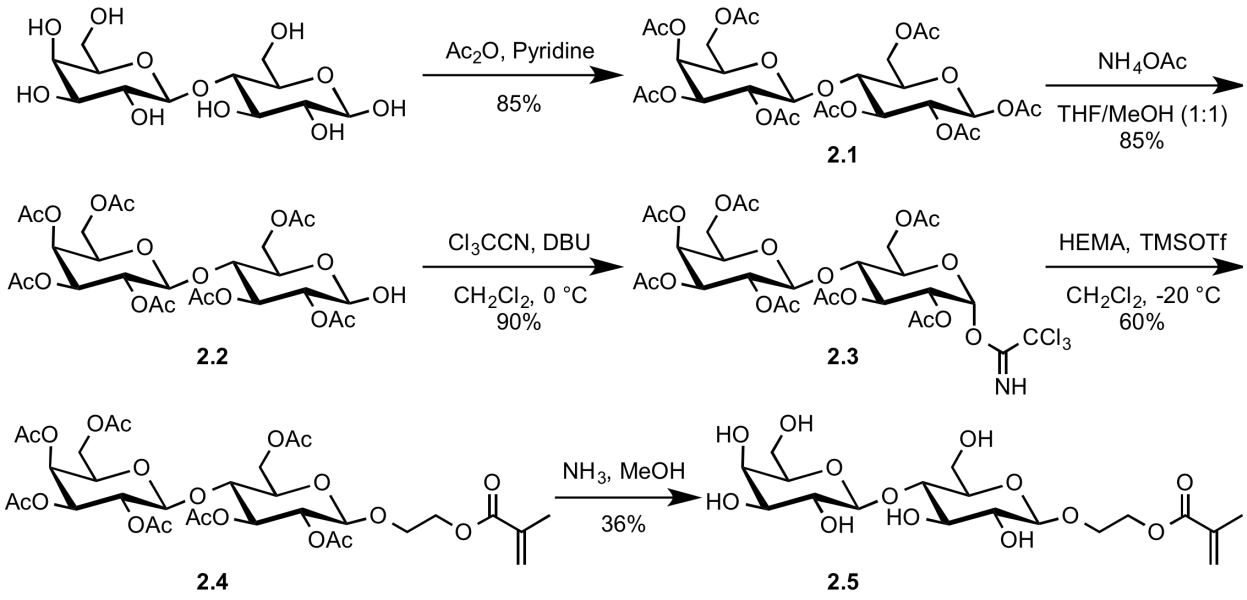
The dissociation constant (K_d) for lactose-galectin interactions is comparable to that with N-acetyllactosamine,[36] and polymers decorated with lactose residues have recently demonstrated the propensity to form complexes with galectins.[37-40] For instance, a lactosyl pseudopolyrotaxane has previously been shown to bind to galectin-1 with a 10-fold enhancement on a per saccharide basis.[41] We therefore focused our early efforts on preparing a lactosyl monomer for CRP because of lactose's commercial availability. We first examined the most direct synthetic route to prepare a lactosyl methacrylate monomer. The Lewis acid-catalyzed glycosylation of per-O-acetylated lactose with 2-hydroxyethyl methacrylate under varying temperatures, concentrations, and ultrasonication methods were unsuccessful. And while the lactosyl methacrylate compound was accessed via the glycosyl bromide donor under Koenigs-Knorr conditions, this route was unrealistic for large-scale preparation due to low yields and major side products.

We next designed a synthetic route based on the Schmidt glycosylation reaction. Lactose would first be fully acetylated and then regioselectively deprotected at the anomeric oxygen. The anomeric position would then be further manipulated to the trichloroacetimidate glycosyl donor.

In the final two transformations, Schmidt glycosylation would be performed with 2-hydroxyethyl methacrylate and followed by global deprotection. We anticipated one drawback from this scheme, namely that the ester within the methacrylate may be labile under the standard acetate deprotection conditions in the final step. Nevertheless, we attempted this concise scheme realizing we would likely need to find mild conditions for ester cleavage in the final step.

β-Lactose (9) was fully acetylated by reaction with acetic anhydride and pyridine. After poor yields were observed for the regioselective deacetylation using ethylenediamine and acetic acid, the per-O-acetylated lactose was deacetylated at the anomeric position with the use of mild ammonium acetate in a tetrahydrofuran and methanol mixture (1:1). Treating heptaacetate **2.2** with trichloroacetonitrile and 1,8-diazobicyclo[5.4.0]undec-7-ene (DBU) led to lactosyl imidate **2.3**, the glycosyl donor. Glycosylation was performed with 2-hydroxyethyl methacrylate and trimethylsilyl trifluoromethanesulfonate as the Lewis acid promoter. The final step towards lactosyl methacrylate was to selectively deprotect the acetyl groups without cleavage of the ester bond in the methacrylate portion of the target. A common procedure for this deprotection step is treatment with a catalytic amount of sodium methoxide in methanol. This condition was investigated by monitoring the reaction over thirty minutes and analyzing its progress via thin-layer chromatography. It was evident that Zemplén conditions were too harsh and cleaved the methacrylate on a rapid time scale. Therefore, a milder deprotection procedure was pursued. Numerous attempts to preserve the methacrylate group with typical organic bases failed. However, when protected **2.4** was treated with a solution of saturated ammonia in methanol (50%), the desired product was produced in a period of 4 h in low yield.

Scheme 2.1. Synthesis of a lactosyl methacrylate monomer.

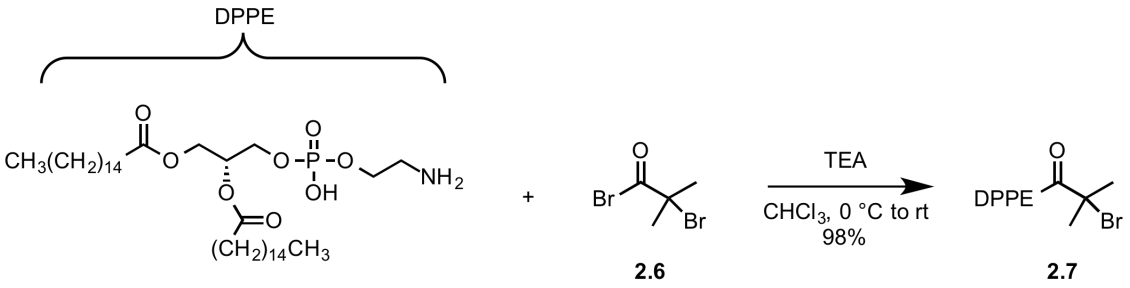


Preparation of lipid-functionalized initiators and chain transfer agents

To examine the CRP of **2.5** with a lipid-functionalized species, we decided to rely on two widely used methods known for their functional group tolerance, atom transfer radical polymerization (ATRP) and reversible addition-fragmentation chain transfer (RAFT)

polymerization. ATRP is typically initiated by an alkyl halide in the presence of a transition metal complex.[42] We sought to construct an ATRP initiator with a lipid functionality conjugated to a tertiary bromide initiator. This strategy would ultimately install the hydrophobic group at one end of the polymer chain. 1,2-Dipalmitoyl-*sn*-glycero-3-phosphoethanolamine (DPPE) was chosen as the phospholipid based on literature precedent for cell surface incorporation.[33] DPPE was reacted with acid bromide **2.6** to form lipid-ATRP initiator **2.7**. It was necessary to perform the reaction with a 1:1 stoichiometry between acyl bromide and the reactive phospholipid since higher equivalents of the electrophile produced inseparable difunctionalized lipid.

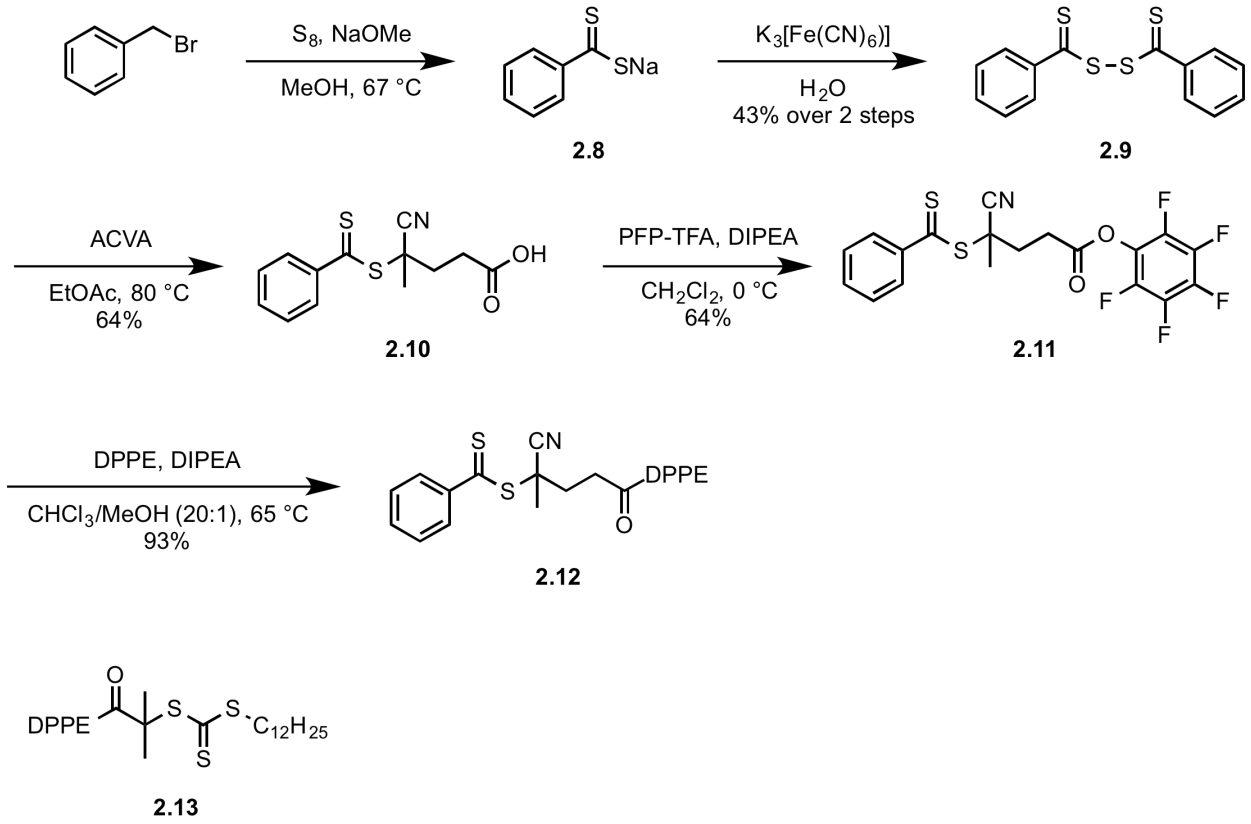
Scheme 2.2. Preparation of lipid-functionalized ATRP initiator.



RAFT polymerization is mediated by a chain transfer agent, usually a thiocarbonylthio compound, that reacts rapidly and reversibly with alkyl radicals to control chain growth.[43] Unlike ATRP, in which the equilibrium between propagating and dormant chains is governed by the transition metal and its coordinated ligands, RAFT is primarily regulated by the judicious choice of chain transfer agent (CTA). Accordingly, we designed both a dithioester and a trithiocarbonate CTA appended with DPPE to facilitate membrane insertion.

Access to the lipid-functionalized dithioester was achieved through a series of reactions to yield a carboxylic acid CTA that could be further elaborated with DPPE. The synthesis commenced by reacting benzyl bromide with elemental sulfur and sodium methoxide to form sodium dithiobenzoate **2.8**. **2.8** was subsequently oxidized using potassium ferricyanide to di(thiobenozyl) disulfide. Disulfide **2.9** was taken forward to the dithiocarboxylic acid by treatment with 4,4'-azobis(4-cyanovaleric acid) (ACVA) under radical conditions. The phospholipid was then attached to the CTA by first converting **2.10** to the pentafluorophenyl (PFP) ester using pentafluorophenyl trifluoroacetate (PFP-TFA) and Hünig's base. **2.11** was finally coupled with DPPE to afford the lipid-functionalized dithioester **2.12** in excellent yield. Lipid-functionalized trithiocarbonate **2.13** was prepared using a similar strategy.

Scheme 2.3. Constructing lipid dithioester CTA **2.12** and trithiocarbonate CTA **2.13**.



CRP of lactosyl methacrylate

Controlled polymerization of lactosyl methacrylate **2.5** was tested using the lipid-ATRP initiator as well as with the two lipid-CTA. No polymerization occurred under any Cu-catalyzed conditions with lipid-initiator **2.7**. As well, short lactosyl oligomers were the only product of RAFT polymerization with trithiocarbonate CTA **2.13**, and these short chains precipitated in the reaction vessel over time. We therefore focused on lipid-functionalized dithioester CTA **2.12**. A mixture of 1,4-dioxane and water (4:1) was found to solubilize both the CTA and the free lactosyl monomer. Our attempts at polymerization at 70 °C with 4,4'-azobis(4-cyanopentanoic acid) as the radical initiator were successful. By following the reaction over time, polymerization appeared to proceed with a linear correlation between molecular weight and monomer conversion (Table 2.1, entry 1). Additionally, low polydispersity indices were obtained for the growing polymer chains at all time points, and the reaction kinetics were defined by a pseudo-first order dependence on lactosyl monomer concentration (Figure 2.2). Taken together, the data indicate that the polymerization was controlled and produced well-defined, lipid-terminated glycopolymers with a degree of polymerization (DP) of 25.

Scheme 2.4. RAFT polymerization of lactosyl methacrylate **2.5**.

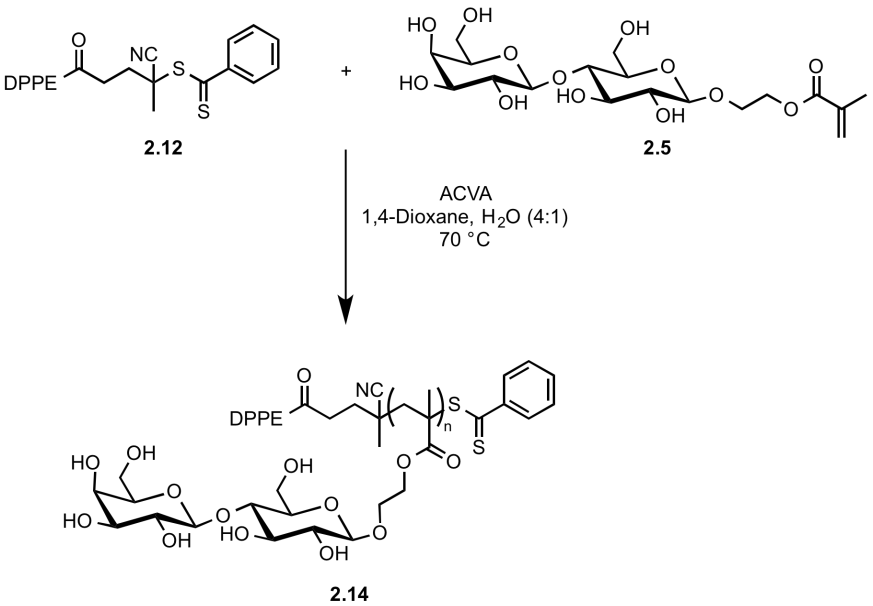


Table 2.1. RAFT polymerization of **2.5** with **2.12** as the RAFT agent. Polymerization was performed in 1,4-dioxane/water mixture (4:1) at 70 °C.

Entry	$[M]_0/[CTA]/[Init]$	Time (min)	M_n (GPC)	Conv. (NMR)	PDI (GPC)
1	50:2:1	80	1,100	0.09	1.25
		251	2,400	0.31	1.23
		436	3,900	0.56	1.22
		1020	6,000	0.95	1.22

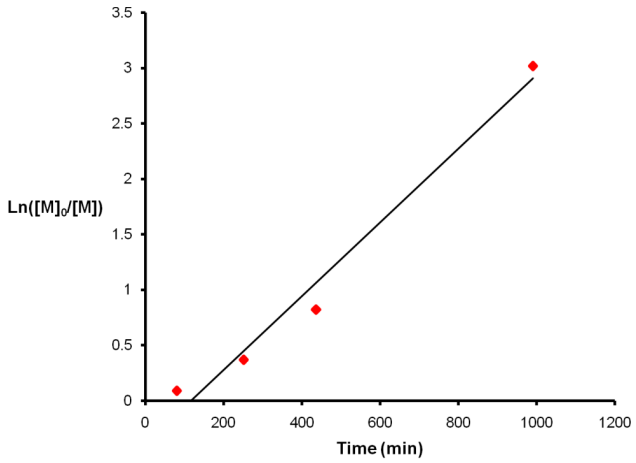
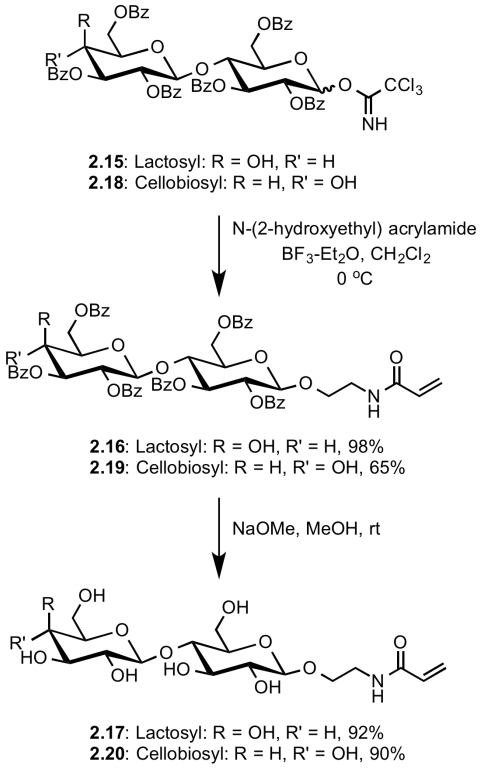


Figure 2.2. Pseudo-first order kinetics for entry **1**.

Synthesis of glycosyl acrylamides

Plagued by low yields of **2.5**, we were unable to prepare sufficient quantities of glycopolymer for further biological studies. To overcome this issue, we transitioned from a methacrylate to a monomer with a more stable amide linkage, i.e. a lactosyl acrylamide. We reasoned that the acrylamide unit would remain intact under the basic conditions necessary for ester cleavage during the final step of monomer synthesis. Starting from the per-O-benzoylated lactosyl trichloroacetimidate **2.15**, we formed the protected acrylamide monomer under Schmidt conditions as before. We next examined whether the large-scale preparation of lactosyl acrylamide was possible. Treating **2.16** with catalytic amounts of sodium methoxide afforded the deprotected glycomonomer **2.17** in excellent yield. Encouraged by these findings, we also tested whether this synthetic strategy could be applied to other glycan structures. Cellobiose, the C-4' epimer of lactose, was chosen as a structurally similar control ligand for our investigations based on its reported lack of binding to galectin.[44] We successfully synthesized cellobiosyl acrylamide **2.20** in short order using the same strategy used for the lactosyl acrylamide. With gram-scale quantities of acrylamide monomers in hand, we turned our attention to their polymerization using RAFT.

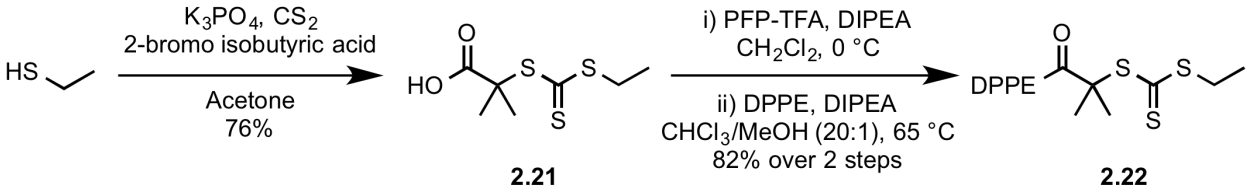
Scheme 2.5. Synthesis of glycosyl acrylamides.



RAFT polymerization of glycosyl acrylamides

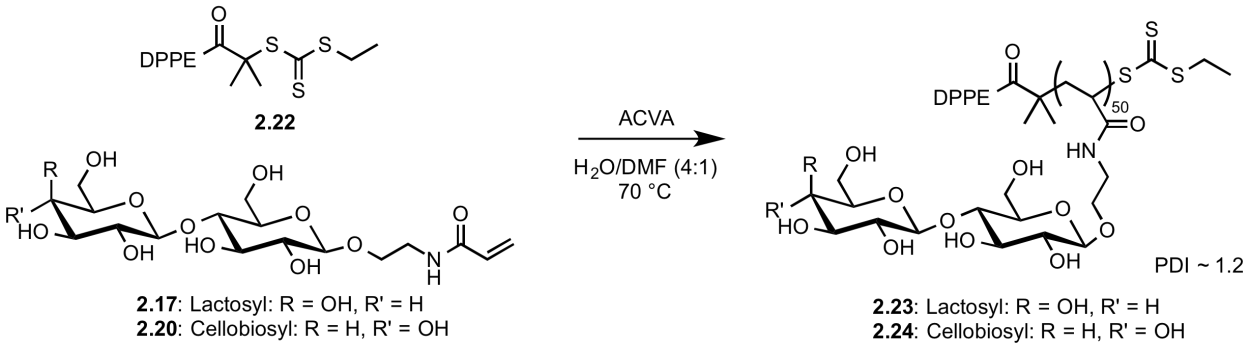
We first attempted the RAFT polymerization of acrylamide **2.17** by employing dithioester CTA **2.12** and trithiocarbonate CTA **2.13**. For both reactions, controlled polymerization proved elusive. In the case of trithiocarbonate CTA **2.13**, we noted precipitation of the growing polymer chain at early time points. This led us to hypothesize that non-covalent association and subsequent aggregation between the two terminal hydrophobic domains impeded continued propagation. As such, we imagined tailoring the trithiocarbonate structure at the 'Z' position to a shorter alkyl chain, minimizing its hydrophobicity. Synthesis of a short alkyl CTA species, i.e. ethyl, commenced by treating ethanethiol with carbon disulfide and 2-bromoisobutyric acid sequentially in a one-pot procedure. Carboxylic acid **2.21** was then transformed to the PFP ester before ligating with DPPE, yielding lipid-functionalized CTA agent **2.22**.

Scheme 2.6. Installation of short ethyl chain at 'Z' position of lipid-functionalized CTA.



The polymerization of lactosyl acrylamide **2.17** in the presence of CTA **2.22** was performed in a mixture of water and *N,N*-dimethylformamide to ensure homogeneity. This reaction indeed generated glycopolymer **2.23** with a low PDI of ~1.2 and in high conversion (>90%). By varying the ratio between monomer and CTA, we were able to construct glycopolymers with predictable DPs of ~25, 50, and 100 (Figure 2.3). We also subjected acrylamide **2.20** to identical RAFT conditions in order to incorporate cellobiosyl residues into the polymers. Again, a well-defined cellobiosyl glycopolymer **2.24** resulted with the use of ethyl CTA **2.22**, which was important for optimal conversion and PDI.

Scheme 2.7. RAFT polymerization of glycosyl acrylamides in the presence of CTA **2.22**.



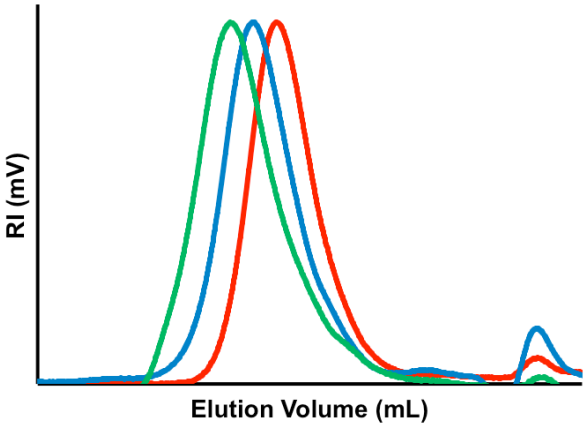


Figure 2.3. Size exclusion chromatography traces of poly(lactosyl acrylamide) lipid-terminated glycopolymers. Polymerizations were performed with $[M]_0$:[CTA] ratios of 100:1 (green), 50:1 (blue), 25:1 (red).

Conclusion

Well-defined glycopolymers, possessing a phospholipid at one end, were synthesized by RAFT polymerization. A lipid-functionalized CTA agent was utilized to control polymerization reactions of glycosyl acrylamide monomers. The final polymers mimic the essential features of membrane-bound glycoprotein ligands of galectins. The lipid domain permits the cell surface incorporation of the glycopolymer into cell membranes, while the polymer portion functions to

interface with galectins native to extracellular environments. A unique feature of this living polymerization system is the ability to form dual end functionalized macromolecules. Glycopolymers **2.23** and **2.24** are well suited for post-polymerization functionalization as the trithiocarbonate remains intact as the polymer end-group. A number of transformations are now at the disposal of chemists to convert the trithiocarbonate moiety to synthetically useful intermediates for further elaboration.[45]

Experimental Procedures

General methods

All of the chemical reagents were of analytical grade, obtained from commercial suppliers, and used without further purification, unless otherwise noted. All of the reactions were performed in a N_2 atmosphere. Liquid reagents were added with a syringe. Dry methanol was purchased from Alfa Aesar in a sealed bottle. All other solvents were purified via packed columns as described by Pangborn *et al*.[46] Unless otherwise noted, sodium sulfate was used as a drying agent and solvent was removed with a rotary evaporator at reduced pressure. Thin layer chromatography was performed on 60 Å glass back silica gel plates and visualized by staining with anisaldehyde or by absorbance of UV light at 245 nm. Flash chromatography was carried out with 60 Å 230-400 mesh silica gel. Unless otherwise noted, ^{1}H, $^{13}C\{^{1}H\}$, ^{19}F and $^{31}P\{^{1}H\}$ NMR spectra were obtained with 300, 400 or 500 MHz Bruker spectrometers. Chemical shifts (δ) are reported in parts per million and standardized against solvent peaks. IR spectra were obtained using thin films of NaCl plates. Polymer M_n was calculated by comparing the area under the peak corresponding to the anomeric proton to that of the methyl protons on the 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine unit (^{1}H NMR). High resolution electrospray ionization (ESI) mass spectra were obtained from the UC Berkeley Mass Spectrometry Facility. UV-Vis spectra were acquired on a Thermo Scientific NanoDrop 2000 Spectrophotometer and used to determine percent modification of fluorophore labeling. Aqueous size exclusion chromatography was performed using a Viscotek TDA 302 SEC fitted with a Shodex SB-803 HQ column and using differential refractive index detection. The mobile phase used was an aqueous solution of $NaNO_3$ (0.1M) at a flow rate of 1.0 mL min^{-1} and at a constant temperature of 40 ^{o}C. Conventional calibration was achieved using narrow PEG standards. Polydispersity indices were determined using the Omnisec for Windows software.

Synthesis

1,2,3,6-Tetra-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-D-galactopyranosyl)-D-glucopyranose (2.1). A mixture of lactose (2.00 g, 5.84 mmol) and acetic anhydride (8.82 mL, 93.5 mmol) in pyridine (20.0 mL) was stirred at room temperature overnight. A 150 mL portion of CH_2Cl_2 was added to the mixture and extracted with three 150 mL portions of 1.0 N HCl. The organic layer was then washed with 150 mL portions of H_2O, satd. $NaHCO_3$, and brine. The resulting solution was dried over $MgSO_4$ and then concentrated. Purification by flash chromatography (1:1 EtOAc/hexane) afforded **2.1** (3.36 g, 85%). All spectroscopic data agreed with the literature.[47] ^{1}H NMR (300 MHz, $CDCl_3$): δ 5.67 (d, J = 8.1 Hz, 1H), 5.34 (dd, J = 3.4, 1.2 Hz, 1H), 5.23 (t, J = 9.0 Hz, 1H), 5.16 – 4.98 (m, 4H), 4.97 – 4.89 (m, 2H), 4.52 – 4.38 (m, 2H), 4.20 – 4.01 (m, 1H), 3.92 – 3.79 (m, 1H), 3.79 – 3.68 (m, 1H), 2.21 – 1.90 (8 x s, 24H).

Lactose heptaacetate (2.2). A flame-dried RB flask was charged with compound **2.1** (1.00 g, 1.47 mmol) and ammonium acetate (227 mg, 2.94 mmol). The mixture was dissolved in 5.00 mL of dry THF/MeOH (1:1) and allowed to react overnight. The mixture was diluted in 100 mL of CH_2Cl_2, extracted with three portions of 100 mL of H_2O, and dried over $NaSO_4$. The crude product was purified by flash chromatography (9:1 → 1:3 hexane/EtOAc) to give **2.2** as a white solid (770 mg, 83%). All spectroscopic data matched the literature.[47] ^{1}H NMR (400 MHz,

$CDCl_3$): δ 5.51 (t, *J* = 9.7 Hz, 1H), 5.35 (dd, *J* = 8.5, 3.5 Hz, 2H), 5.10 (ddd, *J* = 10.2, 7.9, 5.9 Hz, 1H), 4.99 – 4.90 (m, 1H), 4.86 – 4.77 (m, 1H), 4.53 – 4.45 (m, 2H), 4.20 – 4.03 (m, 4H), 3.87 (t, *J* = 6.9 Hz, 1H), 3.80 – 3.72 (m, 1H), 2.21 – 1.93 (7 x s, 21H).

Lactosyl trichloroacetimidate (2.3). Compound **2.2** (1.28 g, 1.91 mmol) was dissolved in dry CH_2Cl_2 (20.0 mL) and placed in an ice bath. Trichloroacetonitrile (959 μL, 9.57 mmol) was added dropwise, and the mixture was stirred for 5 min. 1,8-Diazabicyclo[5.4.0]undec-7-ene (114 μL, 0.765 mmol) was then added dropwise over 2 min and allowed to react for 2 h. The dark orange-brown mixture was then diluted with CH_2Cl_2 and concentrated *in vacuo*. The crude residue was purified by flash chromatography (5:1 → 1:1 hexane/EtOAc with 1% TEA) to yield compound **2.3** as a white solid (1.34 g, 90%). All spectroscopic data agreed with the literature.[48] ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.66 (s, 1H), 6.51 – 6.45 (m, 1H), 5.55 (t, *J* = 9.2 Hz, 1H), 5.38 – 5.32 (m, 1H), 5.09 (ddd, *J* = 30.0, 10.1, 3.2 Hz, 2H), 4.95 (dt, *J* = 8.1, 3.1 Hz, 1H), 4.55 – 4.41 (m, 2H), 4.20 – 4.03 (m, 4H), 3.93 – 3.83 (m, 2H), 2.20 – 1.91 (7 x s, 21H).

Per-O-acetylated lactosyl ethyl methacrylate (2.4). A flame-dried RB flask containing a stir bar and 4 Å molecular sieves was charged with compound **2.3** (1.34 g, 1.71 mmol) and dry CH_2Cl_2 (22.0 mL). To this suspension was added 2-hydroxyethyl methacrylate (250 μL, 2.06 mmol), and the mixture was stirred for 30 min at rt. The RB flask was then transferred to a -20 ^{o}C bath and trifluoromethylsilyl trifluoromethanesulfonate (124 μL, 0.686 mmol) was added dropwise over 2 min. The mixture was allowed to react for 70 min, and trimethylamine was subsequently added to quench the reaction. After the mixture was first warmed to rt and then concentrated, the crude material was purified by flash chromatography (9:1 → 1:1.5 hexane/EtOAc) to afford **2.4** as a white crystalline solid (765 mg, 60%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.05 (s, 1H), 5.53 (s, 1H), 5.32 – 5.23 (m, 1H), 5.13 (t, *J* = 9.2 Hz, 1H), 5.04 (dd, *J* = 10.5, 7.7 Hz, 1H), 4.90 (dd, *J* = 10.4, 3.5 Hz, 1H), 4.84 (t, *J* = 8.7 Hz, 1H), 4.48 (d, *J* = 7.9 Hz, 1H), 4.46 – 4.39 (m, 2H), 4.28 – 4.16 (m, 2H), 4.10 – 3.99 (m, 3H), 3.99 – 3.91 (m, 1H), 3.86 – 3.79 (m, 1H), 3.78 – 3.70 (m, 2H), 3.59 – 3.51 (m, 1H), 2.15 – 1.83 (8 x s, 24H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.32, 170.12, 170.02, 169.74, 169.53, 169.07, 167.07, 136.05, 125.87, 101.07, 100.56, 76.22, 72.76, 72.68, 71.49, 70.96, 70.67, 69.10, 67.48, 66.63, 63.40, 61.97, 60.82, 20.83, 20.79, 20.63, 20.50, 18.26. IR: 2961, 1754, 1635, 1433, 1370, 1223, 1169, 1056 cm^{-1}. ESI-HRMS: Calcd. for $C_{32}H_{44}O_{20}Na^+$ $[M+Na]^+$: 771.2318, found: 771.2291.

Lactose methacrylate (2.5). A dry RB flask was charged with compound **2.4** (500 mg, 0.668 mmol) and a stir bar. Dry MeOH (10.0 mL) was added, and the mixture was stirred for 15 min. A solution of satd. NH_3 in MeOH (10.0 mL) was then added, and the mixture was allowed to react at rt for 4 h. The mixture was diluted in MeOH and concentrated. The crude residue was purified by flash chromatography (1:4 MeOH/$CHCl_3$), diluted in H_2O, and lyophilized to give compound **2.5** as a white solid (110 mg, 36%). ^{1}H NMR (400 MHz, D_2O): δ 6.15 (s, 1H), 5.71 (s, 1H), 4.52 (d, *J* = 8.0 Hz, 1H), 4.42 (d, *J* = 7.7 Hz, 1H), 4.40 – 4.31 (m, 2H), 4.19 – 4.10 (m, 1H), 4.03 – 3.66 (m, 7H), 3.66 – 3.48 (m, 5H), 3.30 (t, *J* = 8.3 Hz, 1H), 1.91 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 168.84, 137.73, 126.44, 105.09, 104.52, 80.57, 77.08, 76.50, 76.39, 74.82, 74.64, 72.55, 70.30, 68.65, 65.25, 62.48, 61.90, 18.41. IR: 3365, 2924, 1745, 1635, 1462, 1378, 1249, 1160, 1056 cm^{-1}. ESI-HRMS: Calcd. for $C_{18}H_{30}O_{13}Na^+$ $[M+Na]^+$: 477.1579, found: 477.1563.

Lipid-functionalized ATRP Initiator (2.7). A flame-dried RB flask was charged with DPPE (100 mg, 0.145 mmol) and $CHCl_3$ (20.0 mL) and stirred at 65 ^oC for 20 min in order to fully dissolve DPPE in $CHCl_3$. The mixture was allowed to cool to rt, and the flask was placed in an ice-bath. Triethylamine (60.4 μL, 0.434 mmol) and 2-bromoisobutyryl bromide (17.9 μL, 0.145 mmol) were added dropwise, and the mixture was stirred overnight and warmed to rt. The reaction mixture was acidified by adding 25 mL of $CHCl_3$ and 40 mL of 0.02 M citrate/0.02 M phosphate buffer (pH 5.5) and stirred for 30 min. The aqueous phase was extracted three times with portions of 40 mL of $CHCl_3$ and discarded. The combined organic phases were dried over anhydrous sodium sulfate, and solvent was removed *in vacuo*. The crude material was purified by flash chromatography (1:19 MeOH/$CHCl_3$ with 1% AcOH) to afford **2.7** as a white solid (120. mg, 98%). ^1H NMR (400 MHz, $CDCl_3$): δ 7.92 (s, 1H), 5.27 – 5.15 (m, 1H), 4.37 (dd, J = 12.0, 3.2 Hz, 1H), 4.15 (dd, J = 12.1, 6.6 Hz, 1H), 3.55 – 3.46 (m, 2H), 2.27 (td, J = 7.6, 5.1 Hz, 4H), 1.94 (s, 6H), 1.64 – 1.50 (m, 4H), 1.36 – 1.12 (m, 48H), 0.86 (t, J = 6.7 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 173.64, 173.27, 172.32, 70.46, 63.85, 62.81, 61.37, 45.69, 34.42, 34.23, 32.16, 32.05, 31.19, 29.84, 29.67, 29.49, 29.28, 25.05, 25.00, 22.81, 14.25, 8.69. ^{31}P NMR (161 MHz, $CDCl_3$): δ 0.40. IR: 3384, 2918, 2850, 1742, 1665, 1535, 1467, 1371, 1223, 1063, 665 cm^{-1}. ESI-HRMS: $C_{41}H_{79}O_9NBrPNa^+$ $[M+Na]^+$: 862.4568, found: 862.4545.

Di(thiobenzoyl) disulfide (2.9). In a 250 mL RB flask, a 25% solution of sodium methoxide in methanol (45.7 mL, 200. mmol), elemental sulfur (6.40 g, 200. mmol), and dry methanol were combined rapidly. Benzyl bromide (11.9 mL, 100. mmol) was then added dropwise over 1 h at rt. After the addition, the mixture was refluxed at 67 ^oC for 14 h. After this time, the reaction mixture was placed in an ice bath. The precipitated salt was removed by filtration and the solvent was removed *in vacuo*. To the residue was added H_2O (100 mL). The solution was filtered a second time and then transferred to a 1 L separatory funnel. The crude sodium dithiobenzoate solution was washed with three portions of 40 mL diethyl ether. Diethyl ether (40 mL) and 1.0 N HCl (100 mL) were added, and dithiobenzoic acid was extracted into the ethereal layer. H_2O (60 mL) and 1.0 N NaOH (120 mL) were added, and sodium dithiobenzoate was extracted to the aqueous layer. This washing process was repeated two more times to finally yield a solution of sodium dithiobenzoate. Due to the propensity of sodium dithiobenzoate to degrade, it was immediately used in the next step. Potassium ferricyanide(III) (6.58 g, 20.0 mmol) was dissolved in H_2O (100.0 mL). The potassium ferricyanide solution was then added dropwise to the sodium dithiobenzoate via an addition funnel over a period of 1 h under vigorous stirring. The pink precipitate was filtered and washed with deionized water until the washings became colorless. The solid was dried *in vacuo*, and the product was recrystallized from ethanol to yield compound **2.9** as a pink solid (6.56 g, 43% over 2 steps). All spectroscopic data matched the literature.[49] ^1H NMR (400 MHz, $CDCl_3$): δ 8.09 (d, J = 7.5 Hz, 1H), 7.62 (t, J = 7.4 Hz, 2H), 7.46 (t, J = 7.8 Hz, 1H).

Dithioester RAFT agent (2.10). In a flame-dried RB flask, compound **2.9** (590. mg, 1.93 mmol) and 4,4′-azobis(4-cyanovaleric acid) (809 mg, 2.89 mmol) were dissolved in 12.0 mL of ethyl acetate. The mixture was then refluxed at 80 ^oC for 20 h and concentrated. The crude material was purified by flash chromatography (2:3 EtOAc/hexane) and recrystallized from benzene to give **2.10** as a red solid (345 mg, 64%). All spectroscopic data matched the literature.[50] ^1H NMR (500 MHz, $CDCl_3$): δ 7.91 (dd, J = 7.5, 1.3 Hz, 1H), 7.58 (t, J = 7.5 Hz, 1H), 7.44 – 7.37 (m, 1H), 2.82 – 2.70 (m, 2H), 2.68 – 2.59 (m, 1H), 2.51 – 2.41 (m, 1H), 1.95 (s, 3H).

PFP-dithioester RAFT agent (2.11). A flame-dried RB flask was charged with compound **2.10** (100. mg, 0.358 mmol) and CH_2Cl_2 (3.00 mL). The mixture was stirred in an ice bath, followed by the dropwise addition of diisopropylethylamine (156 μL, 895 mmol) and pentafluorophenyl trifluoroacetate (74.0 μL, 429 mmol). The mixture was then stirred at 0 ^oC for 1.5 h and concentrated *in vacuo*. The residue was purified by flash chromatography (1:9 → 1:4 EtOAc/hexane) to give **2.11** as a red solid (102 mg, 64%). All spectroscopic data matched the literature.[51] ^1H NMR (400 MHz, $CDCl_3$): δ 7.93 (d, *J* = 8.1 Hz, 1H), 7.59 (d, *J* = 7.4 Hz, 1H), 7.42 (t, *J* = 7.7 Hz, 2H), 3.15 – 2.96 (m, 2H), 2.83 – 2.71 (m, 1H), 2.60-2.52 (m, 1H), 1.99 (s, 3H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -151.69 (d, *J* = 17.3 Hz, 2F), -156.45 (t, *J* = 21.7 Hz, 1F), -161.05 (d, *J* = 19.6 Hz, 2F).

Lipid-functionalized dithioester RAFT agent (2.12). A flame-dried RB flask was charged with compound **2.11** (100. mg, 0.225 mmol), DPPE (155 mg, 0.225 mmol) and a solution of chloroform/methanol (20:1, 5.00 mL). Diisopropylethylamine (196 μL, 1.12 mmol) was added and the mixture was stirred at 65 ^oC. After 5 h, the mixture was allowed to cool to rt and concentrated. The crude solid was purified by flash chromatography (1:19 → 1:4 MeOH/$CHCl_3$ with 1% AcOH), azeotroped with toluene to remove excess acetic acid, and dried *in vacuo* overnight to give compound **2.12** as a red solid (200. mg, 93%). ^1H NMR (400 MHz, $CDCl_3$): δ 10.78 (s, 1H), 7.89 (d, *J* = 7.7 Hz, 2H), 7.54 (t, *J* = 7.4 Hz, 1H), 7.44 (s, 1H), 7.37 (t, *J* = 7.7 Hz, 2H), 5.25 – 5.16 (m, 1H), 4.33 (dd, *J* = 12.1, 3.5 Hz, 1H), 4.14 (dd, *J* = 11.9, 6.5 Hz, 1H), 4.04 – 3.89 (m, 4H), 3.71 (q, *J* = 7.0 Hz, 1H), 3.61 (p, *J* = 6.6 Hz, 1H), 2.68 – 2.52 (m, 3H), 2.50 – 2.37 (m, 1H), 2.36 – 2.20 (m, 4H), 1.92 (s, 3H), 1.66 – 1.46 (m, 4H), 1.42 – 1.14 (m, 48H), 0.86 (t, *J* = 6.6 Hz, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 177.44, 173.75, 173.42, 171.72, 144.71, 133.01, 128.63, 126.78, 118.78, 70.20, 64.56, 63.87, 62.52, 53.58, 46.17, 41.92, 34.36, 34.23, 34.19, 32.03, 31.43, 29.82, 29.78, 29.64, 29.48, 29.26, 25.00, 24.04, 22.80, 21.00, 18.22, 14.23, 11.86. ^{31}P NMR (161 MHz, $CDCl_3$): δ 0.36. IR: 3384, 2918, 2850, 1740, 1653, 1558, 1467, 1378, 1225, 1050, 982, 869 cm^{-1}. ESI-HRMS: $C_{50}H_{85}O_9N_2PS_2Na^+$ $[M+Na]^+$: 975.5326, found: 975.5348.

Lipid-functionalized trithiocarbonate RAFT agent (2.13). A flame-dried RB flask equipped with stir bar was charged with compound PFP-trithiocarbonate RAFT agent[52] (100 mg, 0.188 mmol) and DPPE (130. mg, 0.188 mmol). A solution of dry $CHCl_3$/MeOH (20:1, 5.00 mL) was added, followed by diisopropylethylamine (164 μL, 0.940 mmol). The mixture was stirred at 65 °C for 5 h. The mixture was diluted with chloroform, concentrated, and the residue was purified by flash chromatography (1:19 → 1:9 MeOH/$CHCl_3$ with 1% AcOH) to give **2.13** as a yellow solid (135 mg, 69%). ^1H NMR (500 MHz, $CDCl_3$): δ 5.29 – 5.18 (m, 1H), 4.40 – 4.28 (m, 1H), 4.22 – 4.05 (m, 4H), 3.95 – 3.86 (m, 1H), 3.63 – 3.47 (m, 2H), 2.88 – 2.77 (m, 2H), 2.32 (dt, *J* = 14.7, 7.6 Hz, 4H), 2.23 – 1.85 (m, 12H), 1.35 – 1.17 (m, 66H), 0.88 (t, *J* = 6.6 Hz, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 173.88, 173.56, 173.43, 148.76, 57.35, 37.23, 34.49, 34.30, 32.10, 29.95, 29.86, 29.71, 29.56, 29.47, 29.40, 29.26, 29.19, 27.91, 25.76, 25.15, 25.07, 22.86, 14.29. ^{31}P NMR (161 MHz, $CDCl_3$): δ 0.22. IR: 3368, 2919, 2850, 2360, 1739, 1654, 1522, 1467, 1363, 1243, 1072, 814 cm^{-1}. ESI-HRMS: $C_{54}H_{105}O_9NPS_3^+$ $[M+H]^+$: 1038.6684, found: 1038.6677.

Polymerization of lactose methacrylate (2.14). Polymerizations were carried out using standard Schlenk techniques. Lipid-functionalized dithioester RAFT agent **2.12** (7.08 mg, 0.00735 mmol), lactose methacrylate **2.5** (84.5 mg, 0.45 mmol), 4,4′-azobis(4-cyanovaleric acid)

(1.04 mg, 0.00372 mmol), dioxane-d_8 (640. µl) and D_2O (160. µl) were added to the Schlenk tube. The solution was degassed three times using freeze-pump-thaw cycles. The Schlenk flask was placed in a 70 °C bath to start the polymerization, and aliquots were removed periodically for kinetic analysis. The reactions were stopped by exposure to oxygen. The conversion for the lactosyl monomer was calculated by comparing the integrals of the alkene proton peaks of the monomer (6.15 ppm, 1H) and those of the backbone methyl protons on the glycopolymer (0.50-1.00 ppm, 3H). The solvent was removed and H_2O added to the samples for GPC analysis.

2-*N*-acryloyl-aminoethoxy (2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl)-(1→4)-2,3,6-tri-*O*-benzoyl-β-D-glucopyranoside (2.16). A flame-dried RB flask containing a stir bar and 4 Å molecular sieves was charged with compound **2.16**[53] (0.522 g, 0.429 mmol) and dry CH_2Cl_2 (9.0 mL). To this suspension was added *N*-(2-hydroxyethyl)acrylamide (49.4 µL, 0.472 mmol), and the mixture was stirred for 30 min at rt. The RB flask was then transferred to a 0 °C bath and boron trifluoride diethyl etherate (54.4 µL, 0.429 mmol) was added dropwise over 2 min. The mixture was allowed to react for 90 min, and trimethylamine was subsequently added to quench the reaction. After the mixture was first warmed to rt and then concentrated, the crude material was purified by flash chromatography to afford **2.16** as a white crystalline solid (0.492 g, 98%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.06 – 7.86 (m, 12H), 7.73 (dd, *J* = 8.5, 1.4 Hz, 2H), 7.67 – 7.25 (m, 17H), 7.25 – 7.11 (m, 4H), 6.08 (dd, *J* = 17.0, 1.5 Hz, 1H), 5.94 (t, *J* = 5.6 Hz, 1H), 5.82 (dd, *J* = 10.0, 9.1 Hz, 1H), 5.78 – 5.67 (m, 3H), 5.49 – 5.34 (m, 3H), 4.92 (d, *J* = 7.9 Hz, 1H), 4.73 – 4.62 (m, 2H), 4.48 (dd, *J* = 12.2, 4.2 Hz, 1H), 4.26 (t, *J* = 9.5 Hz, 1H), 3.94 (t, *J* = 6.7 Hz, 1H), 3.89 – 3.80 (m, 2H), 3.80 – 3.63 (m, 3H), 3.58 – 3.49 (m, 1H), 3.42 – 3.30 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 165.99, 165.68, 165.49, 165.47, 165.42, 165.32, 164.89, 133.69, 133.64, 133.62, 133.57, 133.54, 133.42, 133.39, 130.52, 130.10, 129.93, 129.85, 129.79, 129.77, 129.67, 129.58, 129.48, 129.45, 129.08, 128.91, 128.76, 128.74, 128.71, 128.67, 128.40, 128.38, 126.37, 101.42, 101.11, 75.95, 73.33, 72.66, 71.96, 71.83, 71.51, 69.98, 69.46, 67.58, 62.24, 61.14, 39.13, 29.83. HRMS (ESI): calcd. for $C_{66}H_{57}O_{19}NNa^+$ $[M+Na]^+$, 1190.3417; found, 1190.3464.

2-*N*-acryloyl-aminoethoxy (β-D-galactopyranosyl)-(1→4)-β-D-glucopyranoside (2.17). A flame-dried RB flask was charged with compound **2.16** (3.07 g, 2.63 mmol) and a stir bar. Dry MeOH (75 mL) was added, and the mixture was stirred for 15 min. A solution of 0.5 M sodium methoxide in MeOH (2.10 mL) was then added, and the mixture was allowed to react overnight at rt until the product had precipitated. Compound **2.17** was isolated by vacuum filtration as a white solid (1.06 g, 92%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 6.25 (dd, *J* = 17.1, 9.8 Hz, 1H), 6.17 (d, *J* = 17.0 Hz, 1H), 5.74 (d, *J* = 9.9 Hz, 1H), 4.44 (dd, *J* = 22.6, 7.9 Hz, 2H), 4.03 – 3.86 (m, 3H), 3.85 – 3.66 (m, 5H), 3.66 – 3.41 (m, 7H), 3.30 (t, *J* = 7.6 Hz, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 171.11, 132.29, 129.91, 105.37, 104.70, 80.74, 77.79, 77.20, 76.72, 75.21, 74.94, 73.38, 70.98, 70.90, 63.48, 62.46, 41.84. HRMS (ESI): calcd. for $C_{17}H_{29}O_{12}NNa^+$ $[M+Na]^+$, 462.1582; found, 462.1591.

2-*N*-acryloyl-aminoethoxy (2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl)-(1→4)-2,3,6-tri-*O*-benzoyl-β-D-glucopyranoside (2.19). A flame-dried RB flask containing a stir bar and 4 Å molecular sieves was charged with compound **2.18**[54] (3.75 g, 3.09 mmol) and dry CH_2Cl_2 (60.0 mL). To this suspension was added *N*-(2-hydroxyethyl) acrylamide (390. mg, 3.39 mmol), and the mixture was stirred for 30 min at rt. The RB flask was then transferred to a 0 °C bath and boron trifluoride diethyl etherate (390. µL, 3.09 mmol) was added dropwise over 2 min. The

mixture was allowed to react for 90 min, and trimethylamine was subsequently added to quench the reaction. After the mixture was first warmed to rt and then concentrated, the crude material was purified by flash chromatography to afford **2.19** as a white crystalline solid (2.33 g, 65%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.99 – 7.91 (m, 10H), 7.80 – 7.70 (m, 4H), 7.60 – 7.37 (m, 13H), 7.31 – 7.20 (m, 8H), 6.09 (dd, *J* = 16.9, 1.6 Hz, 1H), 5.90 (t, *J* = 5.7 Hz, 1H), 5.86 – 5.63 (m, 3H), 5.53 (dd, *J* = 9.9, 7.8 Hz, 1H), 5.45 – 5.33 (m, 3H), 4.97 (d, *J* = 7.9 Hz, 1H), 4.71 – 4.62 (m, 2H), 4.45 (dd, *J* = 12.1, 4.4 Hz, 1H), 4.23 (t, *J* = 9.5 Hz, 1H), 4.06 (dd, *J* = 11.6, 2.7 Hz, 1H), 3.88 – 3.77 (m, 4H), 3.69 – 3.60 (m, 1H), 3.58 – 3.47 (m, 1H), 3.39 – 3.29 (m, 1H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 165.95, 165.78, 165.73, 165.49, 165.45, 165.41, 165.10, 164.84, 133.60, 133.54, 133.44, 133.37, 130.53, 129.90, 129.86, 129.82, 129.74, 129.71, 129.54, 129.53, 129.40, 129.02, 128.69, 128.66, 128.64, 128.62, 128.57, 128.45, 128.38, 126.36, 101.29, 100.98, 76.36, 73.35, 72.86, 72.53, 72.50, 72.12, 72.02, 69.46, 62.76, 62.21, 39.13, 29.82, -1.72. HRMS (ESI): calcd. for $C_{66}H_{57}O_{19}NNa^+$ $[M+Na]^+$, 1190.3417; found, 1190.3469.

2-*N*-acryloyl-aminoethoxy (β-D-glucopyranosyl)-(1→4)-β-D-glucopyranoside (2.20). A flame-dried dry RB flask was charged with compound **2.19** (2.24 g, 1.92 mmol) and a stir bar. Dry MeOH (55 mL) was added, and the mixture was stirred for 15 min. A solution of 0.5 M sodium methoxide in MeOH (1.53 mL) was then added, and the mixture was allowed to react overnight at rt. The mixture was neutralized using DOWEX 50WX8-100 resin, filtered and concentrated to give compound **2.20** as a white solid (761 mg, 90%). ^{1}H NMR (400 MHz, D_2O): δ 6.24 (dd, *J* = 17.2, 9.8 Hz, 1H), 6.15 (d, *J* = 17.0 Hz, 1H), 5.72 (d, *J* = 10.0 Hz, 1H), 4.45 (d, *J* = 7.1 Hz, 2H), 4.02 – 3.83 (m, 3H), 3.82 – 3.64 (m, 4H), 3.64 – 3.51 (m, 4H), 3.51 – 3.23 (m, 5H). ^{13}C NMR (100 MHz, D_2O) δ 168.66, 129.81, 127.43, 102.53, 102.21, 78.54, 75.93, 75.43, 74.71, 74.15, 73.10, 72.81, 69.40, 68.43, 60.51, 59.93, 39.36. HRMS (ESI): calcd. for $C_{17}H_{29}O_{12}NNa^+$ $[M+Na]^+$, 462.1582; found, 462.1591.

***S*-Ethyl-*S'*-(α,α'-dimethyl-α''-acetic acid)trithiocarbonate (2.21).** A flame-dried RB flask was charged with K_3PO_4 (1.00 g, 4.71 mmol) and dry acetone (22 mL). Ethanethiol (0.350 mL, 4.71 mmol) was added and the mixture was stirred for 10 min. Carbon disulfide (0.612 mL, 10.1 mmol) was then added dropwise and the mixture was stirred for 10 min. After the solution turned bright yellow, 2-bromo isobutyric acid (0.787 g, 4.71 mmol) was added and the mixture was stirred overnight. The mixture was concentrated and the crude material was dissolved in CH_2Cl_2 (25 mL). The organic layer was washed with 1 N HCl (25 mL), water (25 mL), and brine (25 mL), dried with sodium sulfate, and concentrated. The residue was purified by flash chromatography to afford **2.21** as a yellow solid (803 mg, 76%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.43 (s, 1H), 3.29 (q, *J* = 7.4 Hz, 2H), 1.72 (s, 6H), 1.33 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 220.67, 179.42, 55.69, 31.42, 25.32, 13.02. HRMS (ESI): calcd. for $C_7H_{11}O_2S_3^-$ $[M-H]^-$, 222.9927; found, 222.9931.

Lipid-functionalized trithiocarbonate chain transfer agent (2.22). A flame-dried RB flask was charged with compound **2.21** (161 mg, 0.718 mmol) and dry CH_2Cl_2 (8 mL). The mixture was stirred in an ice bath, followed by the dropwise addition of *N*,*N*-diisopropylethylamine (312 μL, 1.80 mmol) and pentafluorophenyl trifluoroacetate (150. μL, 0.861 mmol). The mixture was then stirred at 0 ^oC for 2 h and concentrated *in vacuo*. The crude material was dissolved in hexanes and filtered through a silica plug. A flame-dried RB flask equipped with a stir bar was charged with the filtered PFP ester and 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine (491

mg, 0.710 mmol). A solution of dry $CHCl_3$/MeOH (20:1, 7.00 mL) was added, followed by diisopropylethylamine (520 μL, 2.96 mmol). The mixture was stirred overnight at 65 °C, allowed to cool to rt, and concentrated. The crude solid was purified by flash chromatography, azeotroped with toluene to remove excess acetic acid, and dried *in vacuo* overnight to give compound **2.22** as a yellow solid (0.526 g, 82%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 11.15 (s, 1H), 7.46 – 7.37 (m, 1H), 5.19 (s, 1H), 4.35 (d, *J* = 11.5 Hz, 2H), 4.24 (t, *J* = 5.8 Hz, 1H), 4.12 (dd, *J* = 11.7, 6.7 Hz, 2H), 4.00 – 3.87 (m, 4H), 3.65 – 3.51 (m, 2H), 3.50 – 3.40 (m, 2H), 3.25 (q, *J* = 7.3 Hz, 2H), 3.06 – 2.95 (m, 2H), 2.26 (dd, *J* = 11.9, 6.9 Hz, 4H), 1.99 (s, 1H), 1.73 – 1.62 (m, 6H), 1.61 – 1.50 (m, 5H), 1.45 – 1.15 (m, 42H), 0.85 (t, *J* = 6.4 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 220.51, 173.55, 173.14, 172.73, 70.40, 70.32, 64.29, 64.26, 63.76, 63.73, 62.70, 57.29, 53.24, 41.72, 41.37, 41.32, 34.36, 34.19, 32.01, 31.31, 29.80, 29.75, 29.62, 29.45, 29.42, 29.24, 25.82, 25.01, 24.96, 22.78, 18.64, 17.47, 14.22, 12.90, 11.95. ^{31}P NMR (160 MHz, $CDCl_3$): δ -1.57 (s). HRMS (ESI): calcd. for $C_{44}H_{83}O_9NPS_3^-$ $[M-H]^-$, 896.4973; found, 896.4975.

RAFT polymerization of glycosyl acrylamides. Polymerizations were carried out using standard Schlenk techniques. Lipid-functionalized CTA **2.22** (4.09 mg, 4.55 μmol), lactosyl ethyl acrylamide **2.17** or cellobiosyl ethyl acrylamide **2.20** (0.100 g, 228 μmol), 4,4′-azobis(4-cyanovaleric acid) (319 μg, 1.14 μmol), DMF (200 μl) and H_2O (800 μl) were added to the Schlenk tube. The solution was degassed by sparging with N_2 for 30 min. The Schlenk flask was placed in a 70 ^oC bath to start the polymerization. After 16 h, the reactions were stopped by exposure to oxygen. The conversion for the glycosyl monomer was calculated by comparing the integrals of the alkene proton peaks of the monomer (5.7 ppm, 1H) and those of the backbone methylene protons on the glycopolymer (1.87- 1.00 ppm, 2H). Glycopolymers were purified by dialysis against H_2O for 72 h and lyophilized.

Lactosyl glycopolymer (2.23). According to the above procedures, polymer **2.23** was obtained as a light yellow solid (87 mg, 93%). Conversion (NMR) = 0.95. M_n (NMR) = 20670. PDI (SEC) = 1.21. ^{1}H NMR (500 MHz, D_2O): δ 4.67 – 4.25 (m, 45H), 4.23 – 2.94 (m, 354H), 2.48 – 1.87 (m, 24H), 1.87 – 1.36 (m, 34H), 1.36 – 0.97 (m, 23H), 0.83 (t, *J* = 6.5 Hz, 3H).

Cellobiosyl glycopolymer (2.24). Polymer **2.24** was prepared using the above procedure to yield a light yellow solid (91 mg, 89%). Conversion (NMR) = 0.98. M_n (NMR) = 22490. PDI (SEC) = 1.22. ^{1}H NMR (500 MHz, D_2O): δ 4.60 – 4.36 (m, 49H), 4.26 – 3.04 (m, 411H), 2.51 – 1.87 (m, 27H), 1.87 – 1.36 (m, 40H), 1.36 – 1.00 (m, 22H), 0.93 – 0.76 (m, 3H).

References

(1) Gabius, H.; Siebert, H.; André, S.; Jiménez-Barbero, J.; Rüdiger, H. *ChemBioChem* **2004**, *5*, 740-764.

(2) Park, S.; Lee, M.; Shin, I. *Chem. Soc. Rev.* **2008**, *37*, 1579-1591.

(3) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357-2364.

(4) Lanctot, P. M.; Gage, F. H.; Varki, A. P. *Curr. Opin. Chem. Biol.* **2007**, *11*, 373-380.

(5) Sarter, K.; *et al*. *Autoimmunity* **2007**, *40*, 345-348.

(6) Matani, P.; Sharow, M.; Tiemeyer, M. *Front. Biosci.* **2007**, *12*, 3852-3879.

(7) Gagneux, M.; Varki, A. *Glycobiology* **1999**, *9*, 747-755.

(8) Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discovery* **2005**, *4*, 477-488.

(9) Mossman, K. D.; et al. *Science* **2005**, *310*, 1191-1193.

(10) Ohtsubo, K.; Marth, J. D.; *Cell* **2006**, *126*, 855-867.

(11) Nishiwaki, K.; Kubota, Y.; Chigira, Y.; Roy, S. K.; Suzuki, M.; Schvarzstein, M.; Jigami, Y.; Hisamoto, N.; Matsumoto, K. *Nat. Cell Biol.* **2005**, *6*, 31-37.

(12) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364-2369.

(13) Talbot, P.; Shur, B. D.; Miles, D. G. *Biol. Reprod.* **2003**, *68*, 1-9.

(14) Lowe, J. B. *Cell* **2001**, *104*, 809-812.

(15) Rabinovich, G. A.; Toscano, M. A.; Jackson, S. S.; Vasta, G. R. *Curr. Opin. Struct. Biol.* **2007**, *17*, 513-520.

(16) Di Lella, S.; Sundblad, V.; Cerliani, J. P.; Guardia, C. M. A.; Estrin, D. A.; Vasta, G. R.; Rabinovich, G. A. *Biochemistry* **2011**, *50*, 7842-7857.

(17) Lowe, J. B.; Marth, J. D. *Annu. Rev. Biochem.* **2003**, *72*, 643-691.

(18) Collins, B. E.; Paulson, J. C. *Curr. Opin. Chem. Biol.* **2004**, *8*, 617-625.

(19) Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Curr. Opin. Chem. Biol.* **2000**, *4*, 696-703.

(20) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2754-2794.

(21) Miura, Y. *J. Polym. Sci., Part A: Polym. Chem.* **2007**, *45*, 5031-5036.

(22) Spain, S. G.; Cameron, N. R. *Polym. Chem.* **2011**, *2*, 60.

(23) Ladmiral, V.; Melia, E.; Haddleton, D. M. *Eur. Polym. J.* **2004**, *40*, 431-449.

(24) Coullerez, G.; Seeberger, P. H.; Textor, M. *Macromol. Biosci.* **2006**, *6*, 634-647.

(25) Courtney, A. H.; Puffer, E. B.; Pontrello, J. K.; Yang, Z. Q.; Kiessling, L. L. *Proc. Natl. Acad. Sci.* **2009**, *106*, 2500.

(26) Rele, S. M.; Cui, W.; Wang, L.; Hou, S.; Barr-Zarse, G.; Tatton, D.; Gnanou, Y.; Esko, J. D.; Chaikof, E. L. *J. Am. Chem. Soc.* **2005**, *127*, 10132-10133.

(27) Sigal, G. B.; Mammen, M.; Dahmann, G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996**, *118*, 3789-3800.

(28) Tsuchida, A.; Kobayashi, K.; Matsubara, N.; Muramatsu, T.; Suzuki, T.; Suzuki, Y. *Bioconjugate J.* **1998**, *15*, 1047–1054.

(29) Gestwicki, J. E.; Strong, L. E.; Borchardt, S. L.; Cairo, C. W.; Schnoes, A. M.; Kiessling, L. L. *Bioorg. Med. Chem.* **2001**, *9*, 2387–2393.

(30) Rawat, M.; Gama, C. I.; Matson, J. B.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2008**, *130*, 2959-2961.

(31) Godula, K.; Rabuka, D.; Nam, K. T.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 4973-4976.

(32) Godula, K.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 9963-9965.

(33) Rabuka, D.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 5947-5953.

(34) Kufe, D. W. *Nat. Rev. Cancer* **2009**, *9*, 874–885.

(35) Braunecker, W. A.; Matyjaszewski, K. *Prog. Polym. Sci.* **2007**, *32*, 93-146.

(36) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E.; Yagi, F.; Kasai, K.-ichi *Biochim. Biophys. Acta, Gen. Subj.* **2002**, *1572*, 232-254.

(37) Pohl, N. L.; Kiessling, L. L. *Synthesis* **1999**, *1999*, 1515-1519.

(38) André, S.; Cejas Ortega, P. J.; Perez, M. A.; Roy, R.; Gabius, H.-J. *Glycobiology* **1999**, *9*, 1253 -1261.

(39) Gouin, S. G.; García Fernández, J. M.; Vanquelef, E.; Dupradeau, F.-Y.; Salomonsson, E.; Leffler, H.; Ortega-Muñoz, M.; Nilsson, U. J.; Kovensky, J. *ChemBioChem* **2010**, *11*, 1430-1442.

(40) Vrasidas, I.; Andre, S.; Valentini, P.; Bock, C.; Lensch, M.; Kaltner, H.; Liskamp, R. M. J.; Gabius, H.-J.; Pieters, R. J. *Org. Biomol. Chem.* **2003**, *1*, 803-810.

(41) Nelson, A.; Belitsky, J. M.; Vidal, S.; Joiner, C. S.; Baum, L. G.; Stoddart, J. F. *J. Am. Chem. Soc.* **2004**, *126*, 11914-11922.

(42) Matyjaszewski, K.; Xia, J. *Chem. Rev.* **2001**, *101*, 2921-2990.

(43) Moad, G.; Rizzardo, E.; Thang, S. H *Aust. J. Chem.* **2005**, *58*, 379-410.

(44) Lantéri, M.; Giordanengo, V.; Hiraoka, N.; Fuzibet, J.-G.; Auberger, P.; Fukuda, M.; Baum, L. G.; Lefebvre, J.-C. *Glycobiology* **2003**, *13*, 909-918.

(45) Günay, K. A.; Theato, P.; Klok, H.-A. J. *Polym. Sci. A Polym. Chem.* **2013**, *51*, 1-28.

(46) Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. *Organometallics* **1996**, *15*, 1518-1520.

(47) Khan, R.; Konowicz, P. A.; Gardossi, L.; Matulová, M.; Gennaro, S. *Aust. J. Chem.* **1996**, *49*, 293-298.

(48) Tietze, L. F.; Schuster, H. J.; Krewer, B.; Schuberth, I. *J. Med. Chem.* **2009**, *52*, 537-543.

(49) (a) Cheifari, J.; Mayadunne, R. T. A.; Moad, C. L.; Moad, G.; Rizzardo, E.; Postma, A.; Skidmore, M. A.; Thang, S. H. *Macromolecules* **2003**, *36*, 2273. (b) Chong, Y. K.; Krstina, J.; Le, T. P. T.; Moad, G.; Postma, A.; Rizzardo, E.; Thang, S. H. *Macromolecules* **2003**, *36*, 2256.

(50) Aamer, K. A.; Tew, G. N. *J. Polym. Sci. Part A: Polym. Chem.* **2007**, *45*, 5618-5625.

(51) Roth, P. J.; Wiss, K. T.; Zentel, R.; Theato, P. *Macromolecules* **2008**, *41*, 8513-8519.

(52) Godula, K.;Umbel, M. L.; Rabuka, D.; Botyanszki, Z.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2009**, *131*, 10263-10268.

(53) Mei, X.; Heng, L.; Fu, M.; Li, Z.; Ning, J. *Carbohydr. Res.* **2005**, *340*, 2345–2351.

(54) Chwalek, M.; Plé, K.; Voutquenne-Nazabadioko, L. *Chem. Pharm. Bull.* **2004**, *52*, 965–971.

Chapter 3

Investigating Cell Surface Galectin-Mediated Cross-linking on Glycoengineered Cells

Portions of this chapter were modified from: Belardi, B.; O'Donoghue, G. P.; Smith, A. W.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9549 - 9552.

Chapter 3. Investigating Cell Surface Galectin-Mediated Cross-linking on Glycoengineered Cells[a]

Introduction

Many cellular processes are regulated by multimerization of cell surface proteins and lipids.[1] In many systems, biomolecules assemble into higher-order clusters through direct protein-protein interactions. However in some cases, auxiliary proteins provide scaffolding for oligomeric assemblies via recognition of post-translational modifications.[2] The galectins, a family of secreted glycan-binding proteins, are thought to serve such a function by interacting with specific glycan structures covalently bound to cell surface proteins and lipids.[3] Evidence that both the galectins as well as many of their native ligands are multivalent has led to the proposal of a "galectin lattice" model, in which galectins can segregate membrane-associated glycoproteins and glycolipids into discrete microdomains.[4] Galectin-mediated assemblies have been implicated in the regulation of cell signaling, adhesion, migration, and proliferation,[5] and their dysfunctions have been associated with autoimmune disease[6] and cancer.[7] As well, Dennis and coworkers have proposed that galectin lattices can regulate the cell surface half-lives of glycoproteins by retarding their endocytosis.[8]

Despite compelling evidence for the galectins' role in modulating the behaviors of cell surface molecules, galectin-mediated ligand cross-linking has not been directly observed on live cells. A majority of studies addressing the galectins' crosslinking ability have relied on *in vitro* binding assays.[9] In one cell-based study, Nieminen *et al.* demonstrated that galectin-3 exists in a multivalent state, a requirement for cross-linking, on neutrophil and endothelial cell surfaces through Förster resonance energy transfer (FRET) imaging.[10] However, the effects of galectin binding on *ligand* multimerization have not been directly addressed in cell-based systems, a challenge that is exacerbated by the nature of the galectins' endogenous ligands: they comprise a heterogeneous collection of glycoconjugates that share related glycan structures but disparate underlying scaffolds. There is no straightforward means to selectively label such a complex ligand mixture with biophysical probes that would enable studies of their oligomerization. Specific glycoproteins, such as integrins,[11] mucins,[12] the T cell receptor,[6] and EGFR,[8] have been found to bind galectins in biochemical assays. In principle, the influence of galectins on these proteins' cell-surface behavior can be monitored using GFP fusions and fluorescent antibodies. But on live cells, it is likely that only a subset of their heterogeneous glycoforms engage galectins and form oligomers, which complicates analyses focusing only on the protein component of the ligand.

Here we present a new platform for investigating galectin-mediated cross-linking on live cell surfaces utilizing membrane-associated GPs as chemically defined ligands (Figure 3.1). The GPs were designed to possess the following attributes: (1) galectin-binding glycans distributed across the polymer backbone similarly to galectin-binding mucin glycoproteins,[12] (2) a lipid anchor at one end, and (3) a FRET donor or acceptor fluorophore at the other end. The lipid tail enables insertion of the GP into live cell membranes and control of polymer orientation at the cell surface.[13] The fluorescent dyes allow simultaneous monitoring of GP cross-linking by FRET

[a] Geoffrey P. O'Donoghue and Adam W. Smith contributed to work presented in this chapter.

as well as detection of higher-order assemblies by fluorescence correlation spectroscopy (FCS). Using this experimental platform we found direct evidence for the formation of cell-surface ligand clusters in the presence of galectin-1. More broadly, the method should facilitate interrogation of the galectin-lattice model in the physiologically relevant context of cell surfaces.

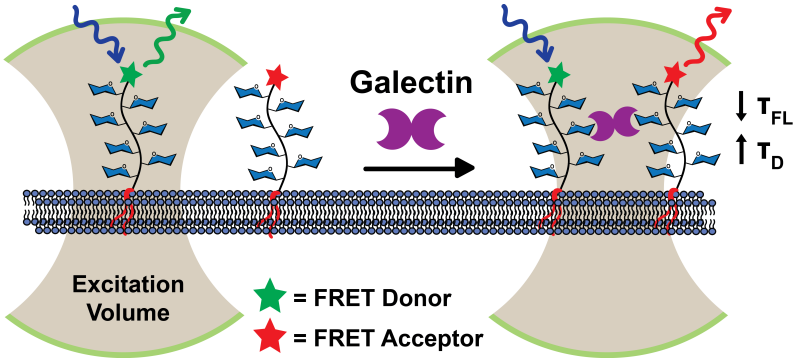


Figure 3.1. An experimental platform for probing galectin-mediated ligand cross-linking on live cell surfaces. Synthetic GPs were adorned with galectin-binding glycans (blue hexagons) and functionalized with a lipid on one end and either a FRET donor or acceptor dye on the other. The GPs were inserted into live cell membranes, and their fluorescence lifetimes (τ_{FL}) and diffusion times (τ_D) were monitored. The galectin-dependent decrease in τ_{FL} and increase in τ_D provided evidence of cell-surface GP cross-linking and oligmerization.

Results and Discussion

Glycoengineering cell surfaces

Well-defined GPs were synthesized by reversible addition-fragmentation chain transfer (RAFT) polymerization (Chapter 2), which enables the dual end-functionalization of the polymer chain. The trithiocarbonate end groups of **2.23** and **2.24** were cleaved with sodium borohydride, and the resulting free sulfhydryl groups were conjugated with maleimide-functionalized Alexa Fluor 488 (donor lactosyl GP **3.1** and donor cellobiosyl GP **3.2**) or Alexa Fluor 555 (acceptor lactosyl GP **3.3** and acceptor cellobiosyl GP **3.4**). These fluorescent lipid-functionalized GPs were used in all subsequent experiments.

We next sought to display polymers **3.1** and **3.2** on live cells for galectin binding studies. Most cell types express endogenous glycoproteins that possess galectin-binding *N*-acetyllactosamine (LacNAc, Galβ1,4GlcNAc) residues.[14] In this initial study, we tried to minimize the impact of endogenous ligands on galectin-GP interactions by choosing a cell line that is deficient in galactosides, the *ldlD* Chinese Hamster Ovary (CHO) cell mutant.[15] *ldlD* CHO cells were incubated with **3.1** or **3.2** for 50 min at rt and imaged using fluorescence microscopy to assess cell surface incorporation. Consistent with previous studies,[13] both GPs produced robust fluorescence localized at the cell membrane as well as within endocytic vesicles (Figure 3.2B and F, and Figures 3.3, 3.4, and 3.5).

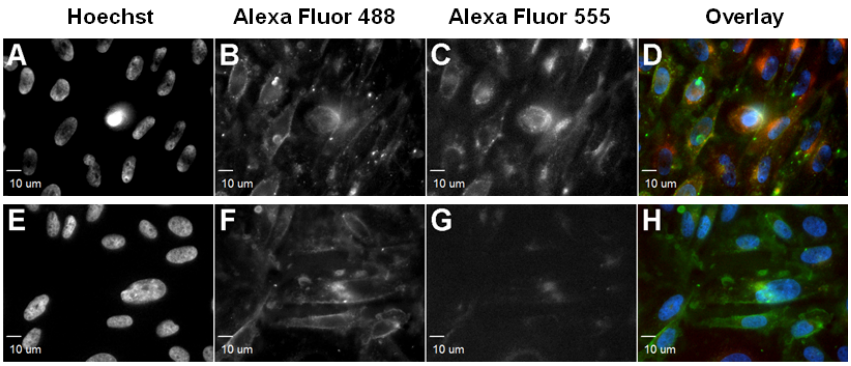


Figure 3.2. Fluorescence microscopy of *ldlD* CHO cells treated with GP **3.1** (A−D) or **3.2** (E−H) followed by fluorescently labeled Galectin-1, Gal-1-555 (D and G). Hoechst 33342 was used to stain the nuclei (A and E). Galectin-1 binding was observed on cells incubated with GP 3.1 (C) but not with GP 3.2 (G).

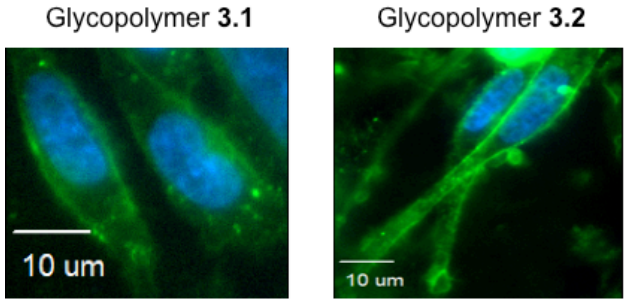


Figure 3.3. Fluorescence microscopy of *ldlD* CHO cells incubated with lipid-terminated fluorescent glycopolymers. The signal (green) demonstrates cell surface incorporation of glycopolymers **3.1** and **3.2**. Hoechst 33342 (blue) was used to stain the nuclei.

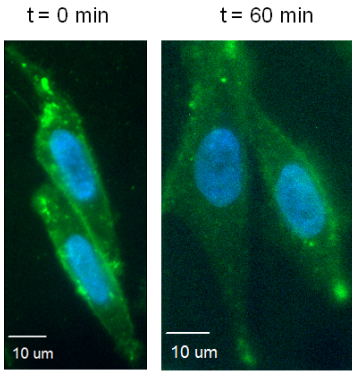


Figure 3.4. Fluorescence microscopy of *ldlD* CHO cells incubated with glycopolymer **3.1** and imaged at different time points. The glycopolymer cell surface signal (green) diminishes over time, consistent with endocytosis of the lipid-terminated glycopolymers. Hoechst 33342 (blue) was used to stain the nuclei.

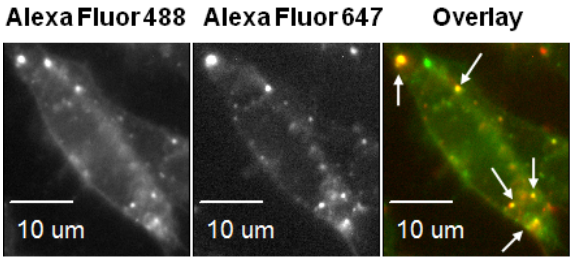


Figure 3.5. Fluorescence microscopy of *ldlD* CHO cells first incubated with glycopolymer **3.1** and then with an early endosome marker, Dextran Alexa Fluor 647. The glycopolymer signal (green) colocalizes with the dextran signal (red), providing evidence that the lipid-functionalized glycopolymers undergo endocytosis and traffic to early endosomes.

Galectin-1 binds to glycopolymer-functionalized cells

For our initial studies we chose galectin-1, a homodimer with two carbohydrate recognition domains (CRDs) at opposite poles of its 3-D structure.[16] The distance between the two CRDs is ~5 nm, well below the Förster radius, R_0, for the FRET pair Alexa Fluor 488 and 555. To explore galectin-1's ability to bind GPs **3.1** and **3.2** when displayed on live cells, galectin-1 was fluorescently labeled with Alexa Fluor 555 (generating Gal-1-555) as previously described.[17] *ldlD* CHO cells previously treated with either **3.1** or **3.2** were incubated with labeled galectin-1, imaged using fluorescence microscopy or analyzed by flow cytometry.

Cells displaying lactosyl GP **3.1** showed significant Gal-1-555 binding (Figure 3.2C), whereas cells displaying cellobiosyl GP **3.2** did not (Figure 3.2G), mirroring the known monomeric ligand preference of galectin-1. Flow cytometry analysis of cells treated similarly gave comparable results (Figure 3.6), although a low amount of Gal-1-555 binding to cell-associated GP **3.2** was observed. It is likely that galectin-1's interaction with cellobiose, though too weak to detect at the monomer level, becomes discernible with multivalent polymers. Overall, these results show that cells deficient in endogenous ligands can be engineered using synthetic GPs to bind galectin-1.

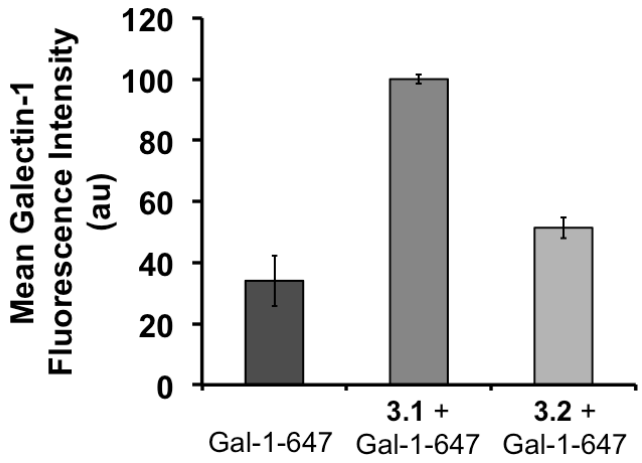


Figure 3.6. Flow cytometry analysis of galectin-1 binding to *ldlD* CHO cells incubated with glycopolymer **3.1**, glycopolymer **3.2**, or no glycopolymer. Error bars indicate the standard deviation.

Monitoring galectin-mediated cross-linking on live cells

The impact of galectin-1 binding on the oligomerization state and mobility of bound GPs was assessed using fluorescence lifetime measurements by time-correlated single-photon counting (TCSPC) and FCS. The principle of the experiment is as follows. GPs **3.1** and **3.2** possess identical backbone and glycan structures but disparate dyes that constitute a FRET pair. The GP pair can be codisplayed on cell membranes in a 1:1 ratio. In the ns regime, fluorescence lifetime measurements can probe for galectin-1-mediated cross-linking since the depletion of excited states of donor **3.1** by proximal acceptor **3.3** ($d < R_0$) decreases the overall fluorescence lifetime (τ_{FL}) of **3.1**.[18] Further, FCS operates by performing an autocorrelation analysis on the fluorescence fluctuations in an ~1 fL excitation volume over many time scales, ns to s.[19] A diffusion time (τ_D) parameter can be extracted from the autocorrelation function, ultimately quantifying the relative mobility of the GPs on the cell surface. The formation of GP clusters would be implied by observed increases in diffusion time.

ldlD CHO cells were first incubated with GPs **3.1** and **3.3** (1:1 ratio), and the time-resolved fluorescence intensity of donor **3.1** was monitored on single live cells with a ps laser pulse at 10 MHz in the presence or absence of unlabeled galectin-1.[20] Fluorescence lifetime and diffusion time were calculated at 10-min intervals. In the absence of galectin-1, donor **3.1'**s fluorescence lifetime increased with time (Figure 3.7). We attribute this phenomenon to endocytosis of the GPs, as evident in our microscopy images. As a consequence, the GPs' density on the cell surface decreases over time, which would reduce the background level of FRET among unclustered GP molecules. Indeed, at lower temperatures at which endocytosis is slower (11− 13 ° C), the fluorescence lifetime of donor **3.1** increased at a slower rate (Figure 3.8).

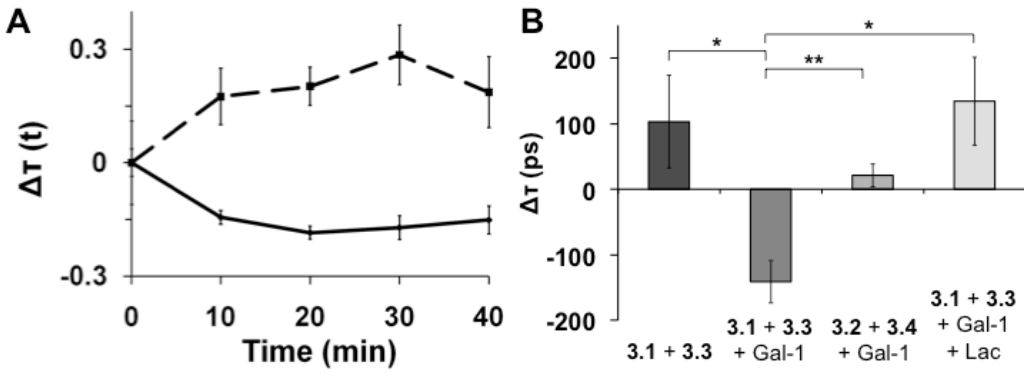


Figure 3.7. Fluorescence lifetime measurements of GPs on *ldlD* CHO cells. (A) Fluorescence lifetime, $\Delta\tau(t)$, measurements of donor 6 in the absence (dotted) or presence (solid) of galectin-1 as a function of time. $\Delta\tau(t)$ represents the difference between $\tau(t)$ and $\tau(0)$. Error bars indicate the standard deviation. (B) The difference in fluorescence lifetime ($\Delta\tau$) between t = 20 and t = 0 min, averaged over six different cells. *P < 0.01; **P < 0.005. Gal-1: unlabeled galectin-1; Lac: 200 mM lactose. Error bars indicate the standard error.

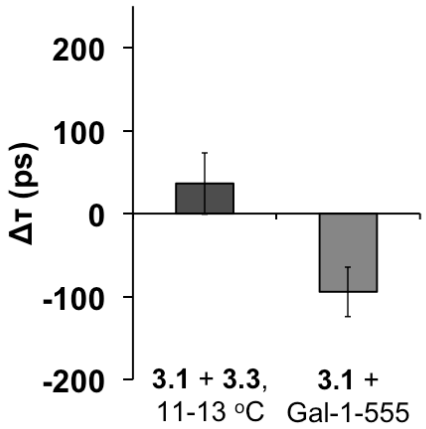


Figure 3.8. The difference in fluorescence lifetime (τ_D) of donor **3.1** on *ldlD* CHO cells between t = 20 and t = 0 min, averaged over six different cells. Error bars indicate standard error.

Next, we monitored the fluorescence lifetime of donor **3.1** after adding galectin-1 to cells labeled with **3.1** and **3.3**. We observed a marked decrease in fluorescence lifetime of donor **3.1** over a period of 40 min (Figure 3.7A). We repeated the experiment on six different cells and observed similar results; the average changes in fluorescence lifetime are shown in Figure 3.7B. These observations suggest that cross-linking of donor **3.1** and acceptor **3.3** by galectin-1 enhanced FRET and, consequently, decreased the fluorescence lifetime of donor **3.1**. We performed a comparable experiment using the cellobiosyl GP FRET pair **3.2** and **3.4**. Despite the presence of galectin-1, the fluorescence lifetime of donor **3.2** increased with time, consistent with endocytosis and little cross-linking. The observed increase was not as dramatic as that observed in the absence of galectin-1, probably reflecting GP **3.2**'s weak interaction with the protein as previously observed by flow cytometry (Figure 3.6). Importantly, the galectin-1-dependent decrease in fluorescence lifetime of GP **3.1** was entirely inhibitable by soluble lactose (200 mM)

(Figure 3.7B). In the presence of this galectin-1 competitor, no significant galectin-1-mediated cross-linking was observed. Additional evidence for a direct interaction between galectin-1 and donor **3.1** was demonstrated through a FRET experiment with fluorescently labeled galectin-1 (Figure 3.8).

Measuring diffusion time of glycopolymers in the presence of galectin-1

Using the same data acquired for fluorescence lifetime measurements, diffusion times were calculated for donor **3.1** over the 40-min time course in the presence or absence of galectin-1. Examples of autocorrelation functions (A and B) and diffusion time values (C) are shown in Figure 3.9. In the absence of galectin-1, diffusion times for donor **3.1** were relatively stable (Figure 3.9C), indicating that the mobility of **3.1** on living cells does not change significantly over time. Notably, the diffusion time for donor **3.1** noticeably increased over time after addition of galectin-1. This dramatic galectin-1-dependent decrease in donor **3.1**'s mobility provides evidence for oligomerization on the cell surface. As before, we averaged the diffusion time values from six different cells and calculated the difference between those values at t = 20 and 0 min (Figure 3.9D). The data confirm that galectin-1 increased the diffusion time of donor **3.1**, which is indicative of reduced mobility and oligomerization of the GP. Free lactose abrogated galectin-1's influence on mobility, consistent with a glycan-binding mechanism of oligomerization. In the presence of 200 mM lactose, the change in diffusion time mirrored that of donor **3.1** in the absence of galectin-1.

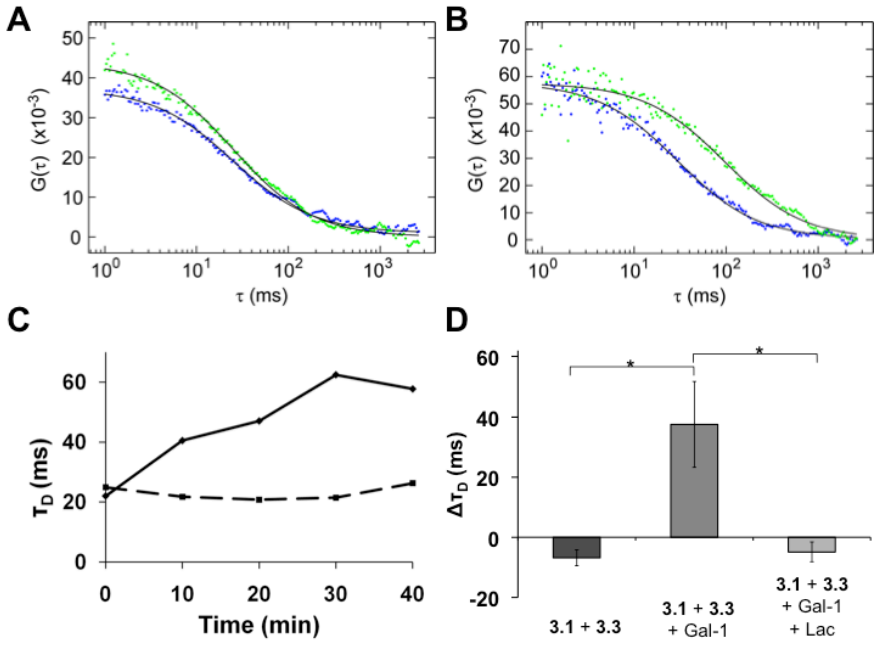


Figure 3.9. FCS of GP **3.1** on *ldlD* CHO cells. Autocorrelation curves for donor **3.1** in the absence (A) or presence (B) of galectin-1 at 0 min (blue) and 20 min (green). (C) Diffusion time, τ_D, for donor **3.1** in the absence (dotted) or presence (solid) of galectin-1. (D) The difference in diffusion time ($\Delta\tau_D$) of donor **3.1** at t = 20 and 0 min averaged over six different cells. *P < 0.05. Gal-1: unlabeled galectin-1; Lac: 200 mM lactose. Error bars indicate the standard error.

Galectin deforms glycoengineered membranes in giant unilamellar vesicles

Having established that galectins assemble glycoconjugates into large, stable cross-linked networks on the cell surface, we wanted to understand the consequences of network formation on membrane curvature. We subsequently inserted GP **3.1** into a simplified membrane system composed of giant unilamellar vesicles (GUVs). GUVs are large liposomal structures that are similar to cell membranes but lack cytosolic components.[21] They are also large enough (10-100 μm) that membrane deformation can be visualized by fluorescent microscopy. We first incubated GP **3.1** with 1-palmitoyl-2-oleoylphosphatidylcholine (POPC)-based GUVs to functionalize the outer leaflet of the membrane with the GP. Using epifluorescence microscopy, we found that the GP uniformly labeled the vesicle without affecting membrane curvature (Figure 3.10A). Upon addition of galectin-3, however, we noted significant membrane bending and deformation of the vesicle structure (Figure 3.10B). Ordered domains, membrane tubes, and endocytic vesicles were all observable and dependent on galectin-mediated cross-linking. These results indicate that galectin clusters can both reinforce and oppose membrane curvature depending on geometric constraints. Taken together, our data suggest that galectins are able to form extended networks on cells surfaces and influence the mechanics and organization of the underlying membrane.

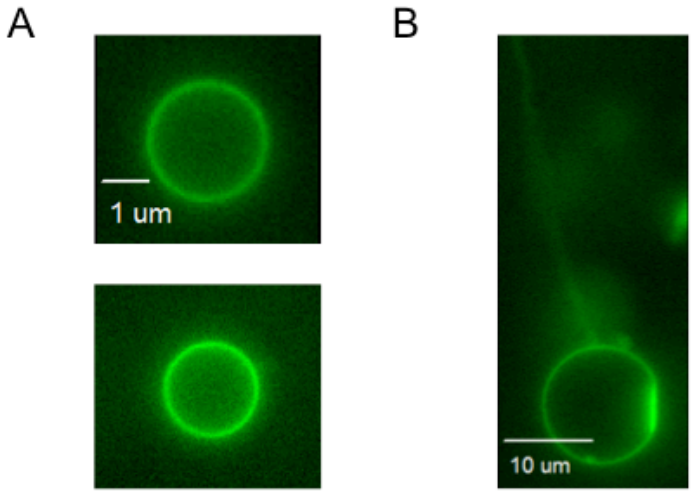


Figure 3.10. Fluorescence microscopy of GUVs first incubated with GP **3.1** (A) and then treated with recombinant galectin-3. Galectin cross-linking of GPs (green) bends POPC membranes.

Conclusion

In conclusion, we have established a new methodology for investigating galectin-glycan interactions on live cell membranes using fluorescently labeled GPs in conjunction with FRET and FCS. To our knowledge, we have also provided the first experimental evidence for galectin-1-mediated cross-linking from the perspective of the bound ligand. Studies are underway addressing the ability of different members of the galectin family to induce such cross-linking. The information should shed light on the dimensions and dynamics of putative galectin lattices on the cell surface as well as the effects of the glycan structure and presentation on galectin recognition and crosslinking. This approach could ultimately provide insight into how the

various galectin members exert different signaling and organizational functions through cell surface interactions. Beyond galectin biology, we envision applications in the study of glycan-receptor interactions between two cells, wherein changes in oligomerization of receptor-bound GPs might reveal the preferred cluster size of the associated glycan-binding protein. As well, the platform can be extended to cell-surface interactions not involving glycans, as the polymers are wholly synthetic and can be adorned with any ligand type.

Experimental Procedures

General methods

All of the chemical reagents were of analytical grade, obtained from commercial suppliers, and used without further purification, unless otherwise noted. All of the reactions were performed in a N_2 atmosphere. Liquid reagents were added with a syringe. Unless otherwise noted, ^{1}H NMR spectra were obtained with 500 MHz Bruker spectrometers. Chemical shifts (δ) are reported in parts per million and standardized against solvent peaks. Polymer M_n was calculated by comparing the area under the peak corresponding to the anomeric proton to that of the methyl protons on the 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine unit. UV-Vis spectra were acquired on a Thermo Scientific NanoDrop 2000 Spectrophotometer and used to determine percent modification of fluorophore labeling. Aqueous size exclusion chromatography was performed using a Viscotek TDA 302 SEC fitted with a Shodex SB-803 HQ column and using differential refractive index detection. The mobile phase used was an aqueous solution of $NaNO_3$ (0.1M) at a flow rate of 1.0 mL min^{-1} and at a constant temperature of 40 ^{o}C. Conventional calibration was achieved using narrow PEG standards. Polydispersity indices were determined using the Omnisec for Windows software.

Alexa Fluor 488 C_5 maleimide, Alexa Fluor 555 C_2 maleimide, Alexa Fluor 647 succinimidyl ester, Dextran Alexa Fluor 647 and Hoechst 33342 nuclear stain were obtained from Invitrogen Life Technologies, Inc. Dulbecco's phosphate-buffered saline pH 7.4 (PBS), Ham's F-12 media, and fetal bovine serum (FBS) were obtained from HyClone Laboratories. Tissue culture plates were obtained from Corning, and glass slides mounted with 8-tissue culture wells were obtained from Nunc. Flow cytometry analysis was performed on a BD FACSCalibur flow cytometer with at least 10^4 live cells for each sample. Cell viability was determined by gating the samples on the basis of forward scatter and side scatter. The average fluorescence intensity was calculated from three, separate replicate samples. Fluorescence microscopy was performed on a Zeiss Axiovert 200M inverted microscope. Images were acquired by using a CoolSNAP HQ charged-coupled device camera. SLIDEBOOK software was used to control the microscope and the camera and for image analysis.

Förster Resonance Energy Transfer (FRET) and Fluorescence Correlation Spectroscopy (FCS) experiments were performed on an inverted microscope (Nikon Eclipse Ti; Technical Instruments, Burlingame, CA). Illumination light for the FRET donor was produced by a 10 MHz, <100 picosecond pulsed diode laser (LDH-P-C-485; PicoQuant GmbH, Berlin, Germany) operating at 488 nm. A single mode fiber was used to reshape the diode laser emission to produce a Gaussian beam profile. Laser power at the sample was set to 1.5 μW by directly modulating the laser power at the laser head using a pulsed diode laser driver (PDL 800-B; PicoQuant GmbH, Berlin, Germany). Several turning mirrors were used to center and shape the focused diode laser beam using both a solution of Alexa Fluor 488 and a DOPC (Avanti Polar Lipids Inc., Alabaster, Al) supported lipid bilayer doped with a small percentage of TopFluor-PS lipids (Avanti Polar Lipids Inc., Alabaster, Al) as a visual indicator of beam quality. Images were recorded using an EM-CCD (iXon 597DU; Andor, Inc., South Windsor, CT). A dichroic beamsplitter (z405/488/561rpc; Chroma Technology Corp., Bellows Falls, VT) reflected light through the objective (Nikon 1.49 N.A. TIRF; Technical Instruments, Burlingame, CA). Fluorescence emission was collected on a 50 μm active area single photon avalanche detector (PDM Series from Micro Photon Devices; Optoelectronic Components, Quebec, Canada).

Bandpass emission filters were FF01-520/44-25 (Semrock Inc., Rochester, NY) and z405/488/561M (Chroma Technology Corp., Bellows Falls, VT). Photon arrival times were measured using an event timer (PicoHarp 300; PicoQuant GmbH, Berlin, Germany) and channel router (PHR 800; PicoQuant GmbH, Berlin, Germany) controlled using the PicoHarp software (PicoQuant GmbH, Berlin, Germany). Photon data were analyzed using custom written MATLAB scripts (The Mathworks, Natick, MA).

Synthesis

Alexa Fluor functionalization of glycopolymers. A RB flask was charged with glycopolymer **2.23** or glycopolymer **2.24** (18.0 mg, 0.871 μmol) and H_2O (10 mL). Sodium borohydride (196 mg, 5.20 mmol) was added to the mixture, and the reaction was stirred for 2 h. Sodium borohyride was removed by dialysis against H_2O for 72 h. The reduced glycopolymer (2.5 mg, 0.13 μmol) was then dissolved in 100 mM Tris buffer (0.4 mL, pH 7.5) and degassed by sparging with N_2 for 15 min. Tris(2-carboxyethyl)phosphine hydrochloride (0.4 mg, 1.3 μmol) was added. After 1 h, Alexa Fluor 488 C_5 maleimide or Alexa Fluor 555 C_2 maleimide (1.0 mg, 1.3 μmol) was added to the reaction mixture. The reaction was stirred overnight at rt in the dark. The fluorescent glycopolymer was purified by dialysis against H_2O for 120 h in the dark and subsequently lyophilized.

Donor lactosyl glycopolymer (3.1). Polymer **3.1** was prepared using the above procedure to yield an orange solid (2.55 mg, 92%, 92% Alexa Fluor modification). M_n (NMR) = 21290. PDI (SEC) = 1.34. See ^{1}H NMR in Appendix.

Donor cellobiosyl glycopolymer (3.2). Polymer **3.2** was prepared using the above procedure to yield an orange solid (2.71 mg, 90%, 99% Alexa Fluor modification). M_n (NMR) = 23050. PDI (SEC) = 1.51. See ^{1}H NMR in Appendix.

Acceptor lactosyl glycopolymer (3.3). Polymer **3.3** was prepared using the above procedure to yield a purple solid (2.44 mg, 86%, 97% Alexa Fluor modification). M_n (NMR) = 21820. PDI (SEC) = 1.35. See ^{1}H NMR in Appendix.

Acceptor cellobiosyl glycopolymer (3.4). Polymer **3.4** was prepared using the above procedure to yield a purple solid (2.55 mg, 83%, 98% Alexa Fluor modification). M_n (NMR) = 23580. PDI (SEC) = 1.52. See ^{1}H NMR in Appendix.

Cell culture

ldlD Chinese hamster ovary (CHO) cells were cultured in Ham's F-12 media supplemented with 10% FBS, 100 units/mL penicillin and 100 μg/mL streptomycin. The cells were maintained at 37 °C and 5% CO_2 in a water-saturated incubator and counted using a hemocytometer. Cell densities were maintained between 1 x 10^5 and 2 x 10^6 cells per mL.

Fluorescence microscopy

Human galectin-1 was recombinantly expressed and purified according to Pace *et al*.[22] Gal-1-555 was prepared from galectin-1 and Alexa Fluor 555 succinimidyl ester following the procedure from Stowell *et al*.[17] *ldlD* CHO cells were seeded at a density of 1 x 10^5 cells per mL on slides mounted with 8-tissue culture wells in 0.3 mL media. After 2 d, the cells were washed with PBS + 1% FBS (2 x 0.5 mL) and then treated with either **6** or **7** (400 μg/mL) in PBS + 1% FBS (0.2 mL). The cells were incubated for 50 min at rt. The cells were then washed with PBS + 1% FBS (3 x 0.4 mL) and incubated with Gal-1-555 (25 μM) in PBS + 1% FBS (0.2 mL) for 40 min at rt and stained with Hoechst 33342 (200 μg/mL) for 5 min. The cells were then washed with PBS + 1% FBS (3 x 0.4 mL) and imaged in PBS + 1% FBS (0.2 mL).

FRET and FCS data collection and analysis

ldlD CHO cells were seeded at a density of 1 x 10^5 cells per mL on slides mounted with 8-tissue culture wells in 0.3 mL media. After 2 d, the cells were washed with PBS + 1% FBS (2 x 0.5 mL) and then treated with either **6** (200 μg/mL) and **8** (200 μg/mL) or **7** (200 μg/mL) and **9** (200 μg/mL) in PBS + 1% FBS (0.2 mL). The cells were incubated for 50 min at rt. The cells were then washed with PBS + 1% FBS (4 x 0.4 mL) and imaged in PBS + 1% FBS (0.2 mL). Cells were first imaged using reflection interference contrast microscopy until three viable cells with suitable size and morphology were found. These cell positions were saved, and each cell was sequentially centered with respect to the focused laser beam. Six 20 s photon streams were collected from the same position inside each cell. This sequence was repeated for the three different cells at 10 min intervals after galectin-1 (25 μM) addition.

Each cell produced photon streams that were then manipulated to yield both a photon arrival histogram and fluorescence correlation curve for each time point before/after galectin-1 addition. The fluorescence lifetime reported for each cell at a particular time point is the decay constant from a single exponential fit of that cells photon arrival histogram. τ_D was chosen as a measure of mobility instead of the diffusion coefficient, because cell membranes often exhibit anomalous (time-dependent) diffusion.[23] τ_D was calculated to be the half-life of the decay of the fluorescence correlation curve.

Galectin-1 binding analysis by flow cytometry

Galectin-1 was labeled with Alexa Fluor 647 succinimidyl ester according to Stowell *et al*.[17] A suspension of 1 x 10^6 *ldlD* CHO cells was pelleted by centrifugation (1300 rpm for 3 min). The supernatants were decanted, and the cells were washed (2 x 0.5 mL) with PBS + 1% FBS (FACS buffer). After resuspension in FACS buffer (0.2 mL), the cells were incubated with **6** or **7** (400 μg/mL) for 50 min at rt. The cells were then pelleted again by centrifugation and washed (2 x 0.5 mL) with FACS buffer. *ldlD* CHO cells treated with either **6** or **7** were subsequently resuspended in FACS buffer (0.2 mL) and incubated with Gal-1-647 (25 μM) for 40 min at rt. The cells were pelleted again, and the supernatant was decanted. The labeled cells were washed (3 x 0.5 mL) and resuspended in FACS buffer (0.4 mL) for analysis by flow cytometry.

Endocytosis analysis by fluorescence microscopy

ldlD CHO cells were seeded at a density of 1 x 10^5 cells per mL on slides mounted with 8-tissue culture wells in 0.3 mL media. After 2 d, the cells were washed with PBS + 1% FBS (2 x 0.5 mL) and then treated with **6** (400 μg/mL) in PBS + 1% FBS (0.2 mL). The cells were incubated for 50 min at rt. The cells were either (1) stained with Hoechst 33342 (200 μg/mL) for 5 min, washed with PBS + 1% FBS (3 x 0.4 mL) and imaged in PBS + 1% FBS (0.2 mL), (2) washed with PBS + 1% FBS (3 x 0.4 mL), incubated in media for 60 min at 37 ^{o}C and 5% CO_2, stained with Hoechst 33342 (200 μg/mL) for 5 min, washed with PBS + 1% FBS (3 x 0.4 mL) and imaged in PBS + 1% FBS (0.2 mL), or (3) stained with Dextran Alexa Fluor 647 (1 mg/mL) and Hoechst 33342 (200 μg/mL) for 5 min at 37 ^{o}C, washed with PBS + 1% FBS (3 x 0.4 mL) and imaged in PBS + 1% FBS (0.2 mL).

Giant Unilamellar Vesicle Functionalization

1-Palmitoyl-2-oleoylphosphatidylcholine (POPC)-based GUVs, prepared by electroformation, were incubated with GP **3.1** (100 μg/mL) overnight at rt in HEPES buffer. GUVs were then isolated by centrifugation (420 x g) and washed (3 x 50 μL) with HEPES. An aliquot of GUV suspension (10 μL) was then removed and imaged by fluorescence microscopy with a flow cell. GP-functionalized GUVs were then treated with recombinant galectin-3 in HEPES for 1.5 h before addition to flow chamber. Membrane curvature was then assessed by fluorescence microscopy in HEPES (10 μL).

References

(1) Groves, J. T.; Kuriyan, J. *Nat. Struct. Mol. Biol.* **2010**, 17, 659.

(2) Ruthenburg, A. J.; Li, H.; Patel, D. J.; Allis, C. D. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 983.

(3) Cooper, D. N. W.; Barondes, S. H. *Glycobiology* **1999**, *9*, 979.

(4) Rabinovich, G. A.; Toscano, M. A.; Jackson, S. S.; Vasta, G. R. *Curr. Opin. Struct. Biol.* **2007**, *17*, 513.

(5) Di Lella, S.; Sundblad, V.; Cerliani, J. P.; Guardia, C. M. A.; Estrin, D. A.; Vasta, G. R.; Rabinovich, G. A. *Biochemistry* **2011**, *50*, 7842.

(6) Demetriou, M.; Granovsky, M.; Quaggin, S.; Dennis, J. W. *Nature* **2001**, *409*, 733.

(7) Liu, F. T.; Rabinovich, G. A. *Nat. Rev. Cancer* **2005**, *5*, 29.

(8) Partridge, E. A.; Le Roy, C.; Di Guglielmo, G. M.; Pawling, J.; Cheung, P.; Granovsky, M.; Nabi, I. R.; Wrana, J. L.; Dennis, J. W. *Science* **2004**, *306*, 120.

(9) (a) Ahmad, N.; Gabius, H.-J.; André, S.; Kaltner, H.; Sabesan, S.; Roy, R.; Liu, B.; Macaluso, F.; Brewer, C. F. *J. Biol. Chem.* **2004**, *279*, 10841. (b) Bourne, Y.; Bolgiano, B.; Liao, D.-I.; Strecker, G.; Cantau, P.; Herzberg, O.; Feizi, T.; Cambillau, C. *Nat. Struct. Biol.* **1994**, *1*, 863.

(10) Nieminen, J.; Kuno, A.; Hirabayashi, J.; Sato, S. *J. Biol. Chem.* **2006**, *282*, 1374.

(11) Moiseeva, E. P.; Williams, B.; Goodall, A. H.; Samani, N. J. *Biochem. Biophys. Res. Commun.* **2003**, *310*, 1010.

(12) Argüeso, P.; Guzman-Aranguez, A.; Mantelli, F.; Cao, Z.; Ricciuto, J.; Panjwani, N. *J. Biol. Chem.* **2009**, *284*, 23037.

(13) Rabuka, D.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 5947.

(14) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. *Essentials of Glycobiology*, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2008.

(15) Kingsley, D. M.; Kozarsky, K. F.; Hobble, L.; Krieger, M. *Cell* **1986**, *44*, 749.

(16) López-Lucendo, M. F.; Solís, D.; André, S.; Hirabayashi, J.; Kasai, K.; Kaltner, H.; Gabius, H.-J.; Romero, A. *J. Mol. Biol.* **2004**, *343*, 957.

(17) Stowell, S. R.; Dias-Baruffi, M.; Penttila, L.; Renkonen, O.; Nyame, A. K.; Cummings, R. D. *Glycobiology* **2003**, *14*, 157.

(18) Levitt, J. A.; Matthews, D. R.; Ameer-Beg, S. M.; Suhling, K. *Curr. Opin. Biotechnol.* **2009**, 20, 28.

(19) Kim, S. A.; Heinze, K. G.; Schwille, P. *Nat. Methods* **2007**, *4*, 963.

(20) *ldlD* CHO cells were washed with 200 mM lactose prior to incubation with GPs in order to remove endogenous galectins.

(21) Dimova, R.; Aranda, S.; Bezlyepkina, N.; Nikolov, V.; Riske, K. A.; Lipowski, R. *J. Phys. Condens. Matter* **2006**, *18*, S1151-76.

(22) Pace, K. E.; Hahn, H. P.; Baum, L. G. In *Methods in Enzymology*; Elsevier, 2003; Vol. 363, pp. 499–518.

(23) Feder, T.; Brustmascher, I.; Slattery, J.; Baird, B.; Webb, W. *Biophys. J.* **1996**, *70*, 2767–2773.

Chapter 4

Imaging the Glycosylation State of Cell Surface Glycoproteins by Two-Photon Fluorescence Lifetime Imaging Microscopy

Portions of this chapter were modified from: Belardi, B.; de la Zerda, A.; Spiciarich, D. R.; Maund, S. L.; Peehl, D. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 14045–14049.

Chapter 4. Imaging the Glycosylation State of Cell Surface Glycoproteins by Two-Photon Fluorescence Lifetime Imaging Microscopy[a]

Introduction

A cell surface protein's glycosylation state can profoundly influence its biological capabilities and can report on the physiological state of the underlying cell.[1-7] Thus, visualization of particular protein glycoforms is an important though challenging goal. Most protein-directed imaging methods (e.g. green fluorescent protein (GFP) tags, fluorescent antibodies) are not sensitive to the glycosylation state of a protein. Our research group previously developed a method or imaging certain glycan structures on native glycoproteins by metabolic labeling with functionalized sugars. However, this glycan-targeted imaging method cannot reveal the identity of proteins to which the labeled glycans are attached.[8-10] Imaging of a specific protein glycoform will require integration of the identifies of both the protein and the glycan.

Other groups have recently made strides towards this goal. Söderberg and coworkers used proximity ligation to detect a specific glycoform of the tumor marker MUC2.[11] More recently, Haga *et al.* used azido sugar labeling of GFP-tagged proteins to image cell surface glycoproteins by Förster resonance energy transfer (FRET) fluorescence microscopy.[12] Due to their reliance on GFP-tagged proteins, however, this method cannot image endogenous glycoproteins or proteins that are not amenable to fluorescent protein fusion. Even so, there are some limitations to a traditional FRET-based technique. The distance between the donor and acceptor fluorophores in a FRET experiment is related to the efficiency of energy transfer and typically precludes the use of two large macromolecules, such as immunoglobulin G (IgG; >10 nm).[13] Another compounding factor for imaging of specific protein glycoforms is the discrepancy between protein copy number and glycan abundance. The difference in abundance between common types of glycans and a specific protein can be orders of magnitude on the cell surface.[14] This large difference in relative number can complicate analyses in imaging applications. For example, in a typical FRET-based experiment, the donor fluorophore is excited, and emission from the acceptor fluorophore is monitored. In the case of high acceptor fluorophore concentration, acceptor bleedthrough can occur causing a false positive FRET signal (Figure 4.1).

[a] Adam de la Zerda, David R. Spiciarich, and Sophia L. Maund contributed to work in this chapter.

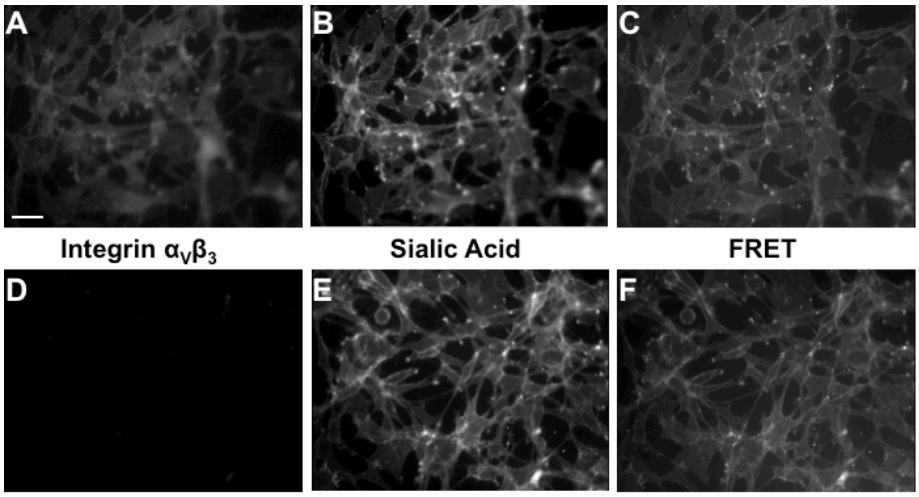


Figure 4.1. Förster Resonance Energy Transfer (FRET) fluorescence microscopy of U87MG cells. Integrin $\alpha_V\beta_3$ was labeled with a donor fluorophore, while the sialic acid residues were labeled with an acceptor fluorophore. FRET is observed even in the absence of donor fluorophore. Scale bar: 50 μm.

Herein, we present a new approach to image endogenous protein glycoforms using a combination of azido sugar labeling and two-photon fluorescence lifetime imaging microscopy (FLIM). We rely on a small (<7 nm) targeting moiety, an antigen-binding fragment (Fab), to introduce the donor fluorophore and locate the protein component. We applied our previously developed glycan labeling strategy to introduce the acceptor fluorophore. In this scheme, cells were first incubated with an azido sugar, peracetylated N-azidoacetylmannosamine (Ac_4ManNAz), which is processed by the cellular machinery and incorporated into glycoproteins as azido sialic acid (SiaNAz). Subsequent bioorthogonal reaction with a cyclooctyne–fluorophore[15,16] conjugate delivers the acceptor fluorophore within a minimal distance (Figure 4.2).

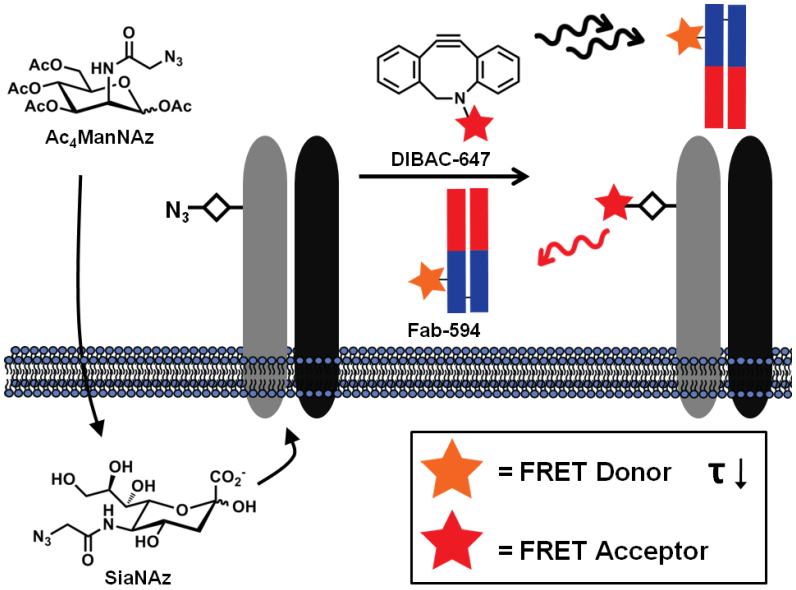


Figure 4.2. A technique for imaging the glycosylation state of glycoproteins through the use of metabolic labeling and two-photon fluorescence lifetime imaging microscopy (FLIM). Cells or tissues are fed the azido sugar, $Ac_4ManNAz$, which is metabolized to SiaNAz, and SiaNAz is then incorporated in sialylated glycoproteins. The azide functionality is reacted with an appropriate bioorthogonal probe bearing an acceptor fluorophore to report on the localization of sialic acid residues. In a subsequent step, the protein scaffold of the glycoprotein is targeted with a donor fluorophore-labeled Fab fragment. The presence of both the FRET donor and acceptor fluorophores on a specific glycoprotein is monitored by changes in the fluorescence lifetime (τ) of the donor after two-photon excitation on a per-pixel basis.

Results and Discussion

SiaNAz residues incorporate into integrin $\alpha_V\beta_3$

A common method for circumventing acceptor bleedthrough is to focus on the donor fluorophore's emission in a FRET experiment. Energy transfer between the donor and acceptor fluorophore results in two major changes to the donor's physical properties. The first is reduction in emission from the donor. Imaging this photon reduction in a population of cells requires normalization by photobleaching of the acceptor to reveal the maximum amount of donor emission, a difficult and tedious task when the field of view contains numerous cells. The other change for the donor fluorophore upon energy transfer is a decrease in fluorescence lifetime.[17] This time-dependent property is advantageous since no further experimentation or sample manipulation is necessary. We sought to utilize the decrease in fluorescence lifetime of the donor fluorophore associated with FRET to monitor the sialylation state of a given glycoprotein through two-photon FLIM.

Overexpression of the integrin $\alpha_V\beta_3$ subtype is observed in variety of cancers and is often correlated with invasiveness due to its pro-angiogenic function.[18] Integrin $\alpha_V\beta_3$ possesses four reported and nine potential N-glycosylation sites on the α subunit and two reported and four potential sites on the β subunit. Recently, several reports have suggested that the glycosylation

state of integrin $\alpha_V\beta_3$ can dramatically alter its migratory and angiogenic functions. In two separate reports, sialylation of integrin $\alpha_V\beta_3$ was found to be required for cell proliferation, migration, and regeneration in wound-healing assays.[19,20] Panjwani and co-workers have shown that $\alpha_V\beta_3$'s glycosylation state also affects VEGF- and bFGF-mediated angiogenesis through an interaction with galectin-3.[21] Motivated by the importance of $\alpha_V\beta_3$'s glycosylation state, we chose to pursue this integrin as a proof-of-concept glycoprotein for imaging its sialylation status.

We first investigated whether endogenous integrin $\alpha_V\beta_3$ is sialylated in the U87MG glioblastoma cell line, which is known to express the heterodimer complex at high levels.[22] $\alpha_V\beta_3$ was immunoprecipitated from U87MG lysate, and the presence of sialic acid on both subunits was assessed by lectin blotting. α2,3- and α2,6-linked sialosides can be distinguished by blotting with *Maackia amurensis* agglutinin (MAA) and *Sambucus nigra* lectins (SNA), respectively. After immunoprecipitation, we observed MAA binding to both integrin subunits, suggesting they are sialylated with α2,3-linked residues (Figure 4.3A).

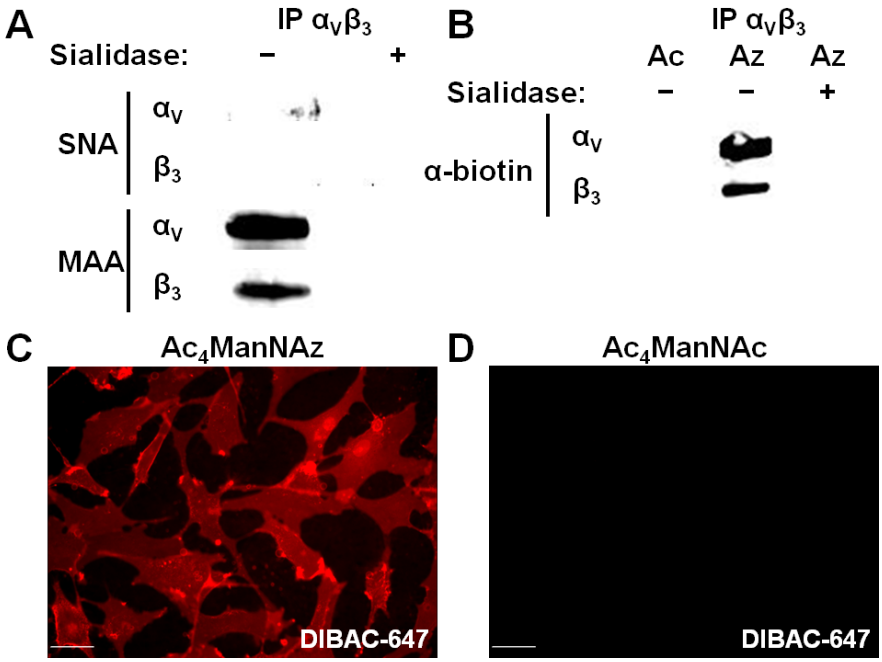


Figure 4.3. Incorporation of SiaNAz into integrin $\alpha_V\beta_3$ and U87MG cell surface glycoconjugates. A) Lectin blot of α and β subunits of immunoprecipitated $\alpha_V\beta_3$ with α2,6 (SNA) and α2,3 (MAA) sialic acid binding lectins in the presence or absence of sialidase. B) Reactivity of $\alpha_V\beta_3$ subunits with phosphine-biotin. Cells were fed either Ac_4ManNAz (Az) or Ac_4ManNAc (Ac) for 3 days and incubated with phosphine-biotin, a probe for the presence of azides, in the presence or absence of sialidase. C,D) Fluorescence microscopy of U87MG cells treated with either Ac_4ManNAz (C) or Ac_4ManNAc (D) and then DIBAC-647 (C,D). Scale bars: 50 μm.

Our FRET scheme is predicated on SiaNAz incorporation into the N-glycans of integrin $\alpha_V\beta_3$ (Figure 1). To confirm this, we supplemented cells with either Ac_4ManNAz or Ac_4ManNAc for 3 days. We then incubated cells with a biotin-conjugated phosphine, a bioorthogonal probe for the azido functionality.[9] Cell lysates were subjected to immunoprecipitation with an anti-

integrin $\alpha_V\beta_3$ antibody, and SiaNAz was detected by western blotting. Azide-dependent labeling was observed for both the alpha and beta chains of integrin $\alpha_V\beta_3$ (Figure 4.3B), suggesting that SiaNAz is present on both subunits of the heterodimer. We also confirmed that the azide-specific signal was indeed due to SiaNAz residues by digesting the lysates with sialidase, whereupon we observed a reduced signal on both integrin subunits (Figure 4.3B).

Preparation and evaluation of Fab-594 and DIBAC-647

We next turned our attention to the selection of the proper FRET pair for the dual labeling strategy. Neuronal cell lines, such as U87MG, are known to exhibit strong flavin- and riboflavin-associated autofluorescence centered at 540-550 nm.[23] To avoid cellular autofluorescence, we focused on donor fluorophores with emission above 600 nm. We ultimately chose the Alexa Fluor 594 (594) and Alexa Fluor 647 (647) fluorophore combination since Alexa Fluor 594's emission maximum is 617 nm, and the Förster radius for the pair is 8.5 nm.

Preparation of the donor Fab fragment conjugate commenced by treating the monoclonal antibody against integrin $\alpha_V\beta_3$ (LM609 clone) with a peptidase that cleaves the full-length IgG between the Fab and Fc segments. The Fab fragment was separated from the Fc segment and unreacted IgG by incubation with protein A resin, and in a subsequent step the lysine residues were reacted with Alexa Fluor 594 *N*-hydroxysuccinimidyl ester to furnish the Fab–594 conjugate. The acceptor cyclooctyne was also synthesized from commercially available reagents, dibenzoazacyclooctyne (DIBAC) amine and Alexa Fluor 647 *N*-hydroxysuccinimidyl ester. The ease of construction of both the donor, Fab-594, and the acceptor, DIBAC-647, targeting moieties underscores the versatility of this method and its applicability to other glycoprotein targets.

After the treatment of U87MG cells with Fab-594, we observed specific cellular labeling of membrane-associated $\alpha_V\beta_3$ that could be inhibited by the presence of full-length anti-$\alpha_V\beta_3$ IgG (Figure 4.4). We also sought to determine if DIBAC-647 was a faithful bioorthogonal probe for azides. Towards this end, we fed cells either Ac_4ManNAz or Ac_4ManNAc for 3 days and then incubated them with DIBAC-647. Cell-surface labeling occurred only in cells administered Ac_4ManNAz (Figure 4.3C,D).

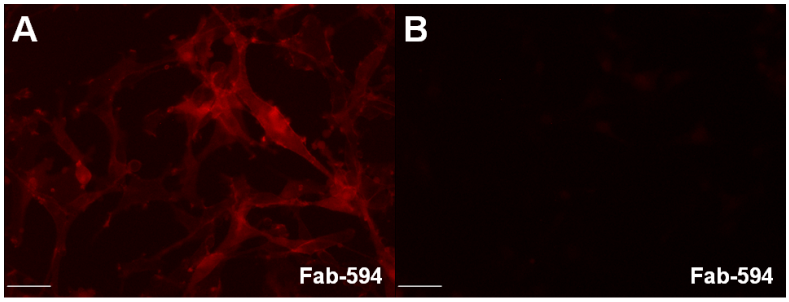


Figure 4.4. Fluorescence microscopy of U87MG cells treated with Ac_4ManNAz and Fab-594 in the absence (A) or presence (B) of anti-$\alpha_V\beta_3$ IgG. Scale bar: 50 μm.

Fluorescence lifetime imaging microscopy of integrin $\alpha_V\beta_3$'s sialylation state

After confirming the labeling of $\alpha_V\beta_3$'s protein backbone with the FRET donor Fab-594 and SiaNAz residues with the FRET acceptor DIBAC–647, we focused on monitoring FRET using two-photon FLIM. The two-photon FLIM FRET technique entails exciting the donor fluorophore using near-IR/IR femtosecond pulses and monitoring the time-dependent fluorescence decay in the nanosecond regime.[17] In the case when an acceptor fluorophore is in close proximity to the donor fluorophore, the excited state of the donor is depleted through energy transfer, which ultimately results in the reduced fluorescence lifetime (τ) of the donor molecule. We observed that Fab-594 had an *in vitro* characteristic τ value of 3.09 ns, which was determined by fitting the fluorescence decay of Fab–594 to a single exponential (Figure 4.5). In FLIM, the fluorescence emission for each pixel of an image is monitored in a time-dependent manner. The emission data are then fitted to extract the τ values, and these are displayed as a heat-map image.[17]

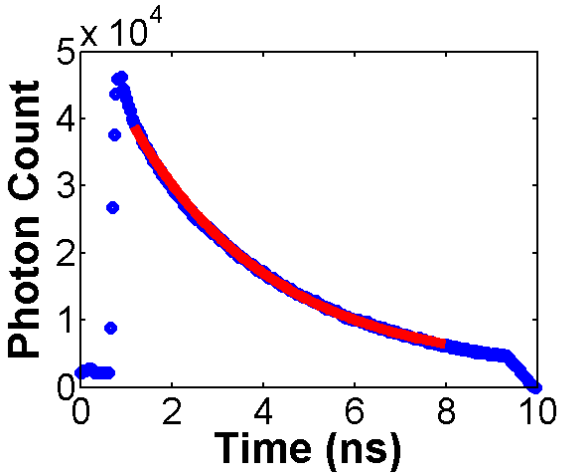


Figure 4.5. Fluorescence decay histogram for Fab-594. The blue dots represent the arrival time of emitted photons with the single exponential fit (red line) overlaid.

FLIM images were obtained by incubating U87MG cells with either Ac_4ManNAz or Ac_4ManNAc for 3 days and then labeling cells with DIBAC-647 and Fab-594, sequentially. Cells that were treated with Ac_4ManNAc had an average τ value of 2.91 ns (Figure 4.6A). The cell-surface regions had a relatively constant τ value, as represented by the deep blue color in the image in Figure 4.6A. On the other hand, in cells treated with Ac_4ManNAz, the average τ value decreased to 2.60 ns (Figure 4.6B), which is consistent with cell-surface FRET[24] between Fab-594 and DIBAC-647; the cell-surface τ values appeared far more heterogeneous than in Figure 4.6A, as evident by green and yellow clusters. Collectively, these data indicate that cell surface integrin $\alpha_V\beta_3$ is sialylated on U87MG cells. We further analyzed the τ values by constructing a histogram of the cell surface τ values for both Ac_4ManNAc and Ac_4ManNAz cell populations (Figure 3C). The decrease in the τ values from cells labeled with only the Fab-594 donor versus cells with both the Fab-594 donor and the DIBAC-647 acceptor was also accompanied by a switch from a unimodal (red) to a multimodal distribution (blue). These changes may reflect variations in the amount of sialylation on individual integrin $\alpha_V\beta_3$ molecules. As well, cells

lacking Fab-594 but labeled with the DIBAC-647 acceptor displayed minimal signal (Figure 4.7A).

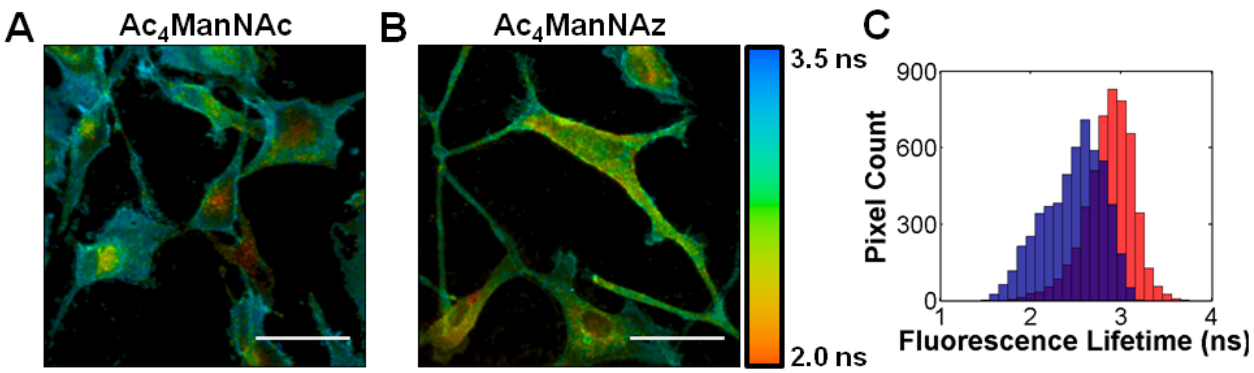


Figure 4.6. FLIM imaging of sialylated integrin $\alpha_V\beta_3$. A,B) Two-photon FLIM images of U87MG cells treated with either Ac_4ManNAc (A) or Ac_4ManNAz (B), followed by DIBAC-647 and Fab-594. FLIM images are the heat maps of τ values determined per pixel (see accompanying legend, right). C) Histogram of cell surface τ values for cells treated with either Ac_4ManNAc (red bars) or Ac_4ManNAz (blue bars) and then incubated with DIBAC-647 and Fab-594. Scale bars: 50 μm.

FRET is integrin- and glycan-dependent

We further sought to evaluate whether the observed FRET arose from SiaNAz residues on integrin $\alpha_V\beta_3$ itself, or alternatively, from nearby sialic acid-modified glycoproteins or glycolipids. We reasoned that by adding an acceptor fluorophore to cell surfaces nonspecifically, while guaranteeing that the acceptor moiety was not bound to our integrin-of-interest, we could differentiate between sialylated integrin-dependent FRET and background FRET. Cells were first incubated with dipalmitoyl phosphatidylethanolamine (DPPE)-647, which incorporates the acceptor into cell membranes uniformly, and then followed with Fab-594 labeling. No decrease in τ values was observed in this case as cell surfaces appeared to maintain a similar hue to that of cells that only possess the Fab-594 fragment (Figure 4.7B,D). Thus, we reasoned that sialic acid residues on nearby proteins or lipids do not influence the FRET signal originating from $\alpha_V\beta_3$ complexes. We also tested whether the observed FRET was a sialylation state-dependent phenomenon. We treated Ac_4ManNAz-fed U87MG cells with a sialidase to cleave SiaNAz residues from cell surface glycoconjugates. The cells were further treated with DIBAC-647 and incubated with Fab-594. Sialidase treatment abolished the decrease in τ (Figure 4.7C,D), indicating that SiaNAz residues on integrin $\alpha_V\beta_3$ were critical for the observed FRET.

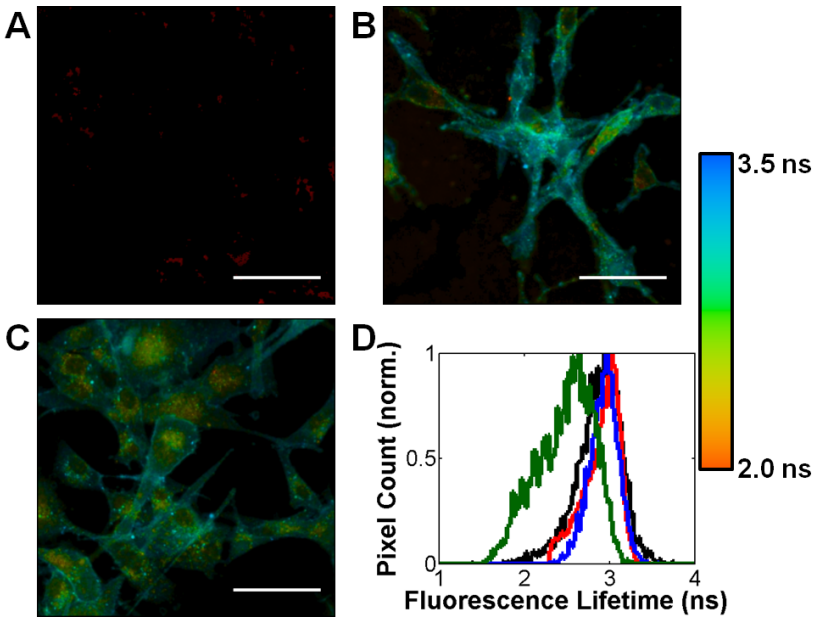


Figure 4.7. FLIM FRET is dependent on the proximity of the acceptor dye to Fab-594 and the target protein's glycosylation state. FLIM images of A) Ac_4ManNAz-fed U87MG cells treated with DIBAC-647 alone, B) U87MG cells modified with DPPE-647 and then incubated with Fab-594, and C) Ac_4ManNAz-fed U87MG cells first treated with sialidase and then incubated with DIBAC-647 and Fab-594. FLIM images are the heat maps of τ values determined per pixel (see accompanying legend, right). D) Histogram of cell surface τ values for cells treated with either Ac_4ManNAc (black line), DPPE-647 (red line), Ac_4ManNAz and sialidase (blue line), or Ac_4ManNAz (green line) and then incubated with DIBAC-647 (black, blue, and green) and Fab-594 (black, red, blue, and green). Scale bars: 50 μm.

Imaging the glycosylation state of integrin $\alpha_V\beta_3$ in human prostate adenocarcinoma tissue slices

A unique attribute of this platform is that it can monitor the glycosylation state of endogeneous glycoproteins as well as image deep within a specimen, owing to the two-photon mode of excitation. We have previously shown that human prostate adenocarinoma tissue slices can metabolically incorporate azido sugars into their glycoproteins.[25] Prostate cancers are known to undergo an upregulation in integrin $\alpha_V\beta_3$,[26] which provided us an opportunity to apply our imaging technique to human tissue slices. Grade 3-4 precision-cut adenocarinoma tissue slices derived from an 8 mm core of a radical prostatectomy were cultured for 3 days in either Ac_4ManNAz or Ac_4ManNAc.We first determined that $\alpha_V\beta_3$ immunoprecipitated from tissue lysate, like $\alpha_V\beta_3$ derived from U87MG lysate, was recognized by MAA (Figure 4.8A) and incorporated SiaNAz residues into its glycans (Figure 4.8B). Using fluorescence microscopy, Fab-594 and DIBAC-647 were found to label prostate tissue slices in a $\alpha_V\beta_3$- and azide-dependent manner (Figures 4.9 and 4.10), respectively. We next applied our FLIM FRET strategy to the tissue slices by first staining with DIBAC-647 and then with Fab-594. Ac_4ManNAc-treated tissue slices displayed the characteristic Fab-594 τ value (Figure 4.11A,D), as evident by the blue color. The Ac_4ManNAz-treated tissue slices showed a decrease in τ values (Figure 4.11B,D) that could be abrogated by treatment with sialidase (Figure 4.11C,D), again consistent with FRET and the presence of sialylated integrin $\alpha_V\beta_3$ embedded within the tissue

slices. These data show that we can probe protein glycosylation with increasing molecular precision.

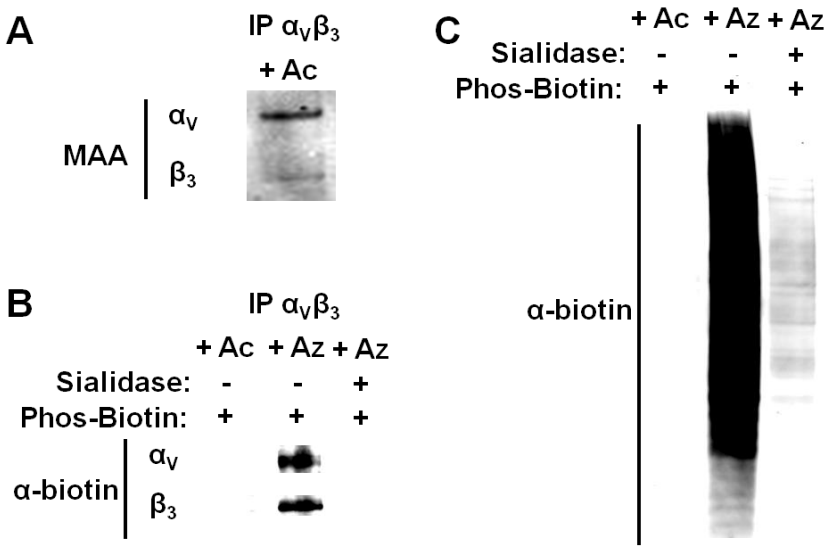


Figure 4.8. (A) Lectin blot of α and β subunits of immunoprecipitated $\alpha_V\beta_3$ from human prostate adenocarcinoma lysate with MAA in the presence or absence of sialidase. Reactivity of $\alpha_V\beta_3$ subunits (B) from human prostate adenocarcinoma lysate (C) with phosphine-biotin. Prostate tissue slices were fed either Ac_4ManNAz (Az) or Ac_4ManNAc (Ac) for 3 days and incubated with phosphine-biotin in the presence or absence of sialidase.

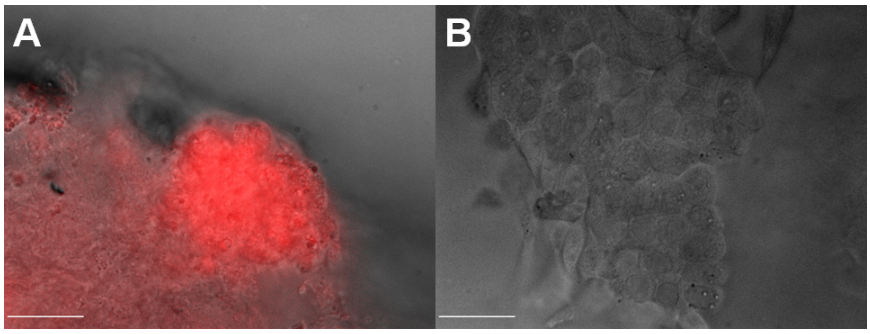


Figure 4.9. Fluorescence microscopy of human prostate adenocarcinoma tissue slices treated with Fab-594 in the absence (A) or presence (B) of anti-$\alpha_V\beta_3$ IgG. Fab-594 fluorescence (red) overlay with brightfield. Scale bar: 50 μm.

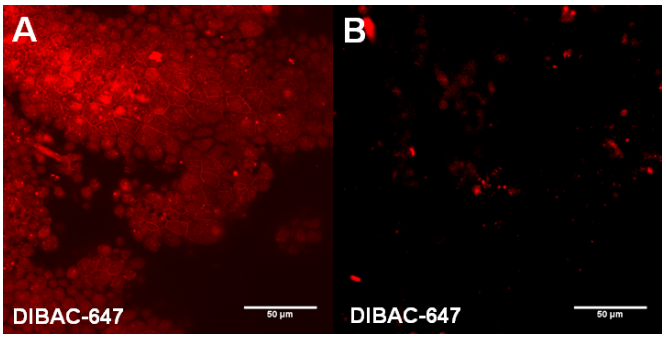


Figure 4.10. Confocal fluorescence microscopy of human prostate adenocarcinoma tissue slices treated with either Ac_4ManNAz (A) or Ac_4ManNAc (B) for 3 days and then incubated with DIBAC-647.

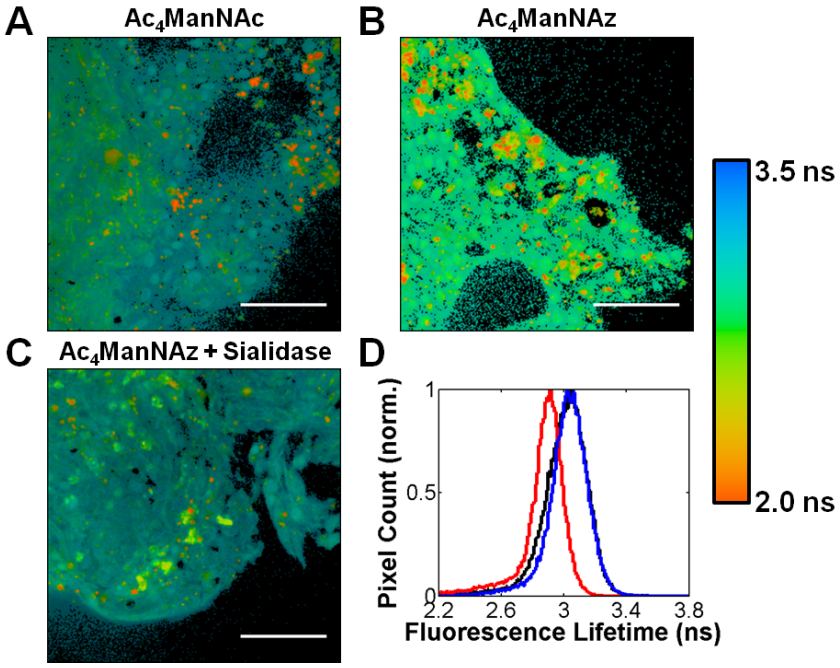


Figure 4.11. FLIM images and lifetime histogram of grade 3-4 human prostate adenocarinoma tissue slices. Tissue slices were incubated with either Ac_4ManNAc (A), Ac_4ManNAz (B), or Ac_4ManNAz and sialidase (C) and then treated with DIBAC-647 and Fab-594. FLIM images are the heat maps of τ values determined per pixel, (see accompanying legend, right). D) Histogram of τ values for tissues treated with either Ac_4ManNAc (black line), Ac_4ManNAz and sialidase (blue line), or Ac_4ManNAz (red line) and then incubated with DIBAC-647 and Fab-594. Scale bars: 50 µm.

Conclusion

In conclusion, we have developed a FLIM FRET-based method for visualizing the glycosylation state of specific glycoproteins. The method overcomes intrinsic limits of classic FRET imaging, which can be undermined by bleedthrough. As well, we can now also image

discrete protein glycoforms in whole tissue slices. The method therefore adds to a growing toolkit for characterizing cell surface glycomes with molecular precision. We plan on utilizing this technique in the future for imaging the cognate cell surface ligands for secreted lectins as well as for diagnostic purposes in monitoring the glycosylation status of specific glycoproteins.

Experimental Procedures

General methods

All of the chemical reagents were of analytical grade, obtained from commercial suppliers, and used without further purification, unless otherwise noted. Alexa Fluor 594 carboxylic acid, succinimidyl ester and Alexa Fluor 647 carboxylic acid, succinimidyl ester were obtained from Life Technologies, Inc. DIBAC-amine was purchased from Click Chemistry Tools. Ac_4ManNAz and Ac_4ManNAc were synthesized as previously described.[27] All of the reactions were performed in a N_2 atmosphere. Liquid reagents were added with a syringe. Unless otherwise noted, ^{1}H and $^{13}C\{^1H\}$ spectra were obtained with either 400 or 500 MHz Bruker spectrometers. Chemical shifts (δ) are reported in parts per million and standardized against solvent peaks. High resolution electrospray ionization (ESI) mass spectra were obtained from the UC Berkeley Mass Spectrometry Facility. UV-Vis spectra were acquired on a Thermo Scientific NanoDrop 2000 Spectrophotometer. Reversed-phase HPLC was performed on a Varian system attached to an absorption detector using a C18 reverse phase column (5 μm, 250 x 4.6 mm, Agilent) for purifications. HPLC solvents were A: H_2O with 0.1% TFA and B: MeCN with 0.1% TFA. Blot fluorescence was measured with a Typhoon 9410 imaging system (Amersham).

Dulbecco's phosphate-buffered saline pH 7.4 (PBS), Dulbecco's modified Eagle's media (DMEM) without phenol red, and fetal bovine serum (FBS) were obtained from HyClone Laboratories. Antibodies and lectins were purchased from EMD Millipore, Jackson ImmunoResearch and Vector Laboratories. Tissue culture plates were obtained from Corning, and glass slides mounted with 4- and 8-tissue culture wells were obtained from Nunc. Fluorescence microscopy was performed on a Zeiss Axiovert 200M inverted microscope equipped with a 20x/0.75 Fluar air lens. Images were acquired by using a CoolSNAP HQ charged-coupled device camera (Roper Scientific). A 175 W Xenon lamp housed in a Sutter DG4 illuminator linked to the microscope by an optical fiber assured shuttering and illumination. SLIDEBOOK software (Intelligent Imaging Innovations, Inc.) was used to control the microscope and the camera and for image analysis.

Time-correlated single photon counting (TCSPC) data sets were acquired on an inverted Zeiss LSM510 NLO microscope with a Plan-Neofluar 40x/1.30 NA oil immersion lens equipped with a Becker & Hickl SPC-830 TCSPC controller. Samples were excited by two-photon, femtosecond 800 nm pulses generated by a 80 MHz Ti:sapphire laser (Coherent MiraSHG). Several turning mirrors were used to center and shape the focused laser beam using a fluorescent slide (Microscopy/Microscopy Education). The non-descanned emission was collected from a back port, filtered through both a 700 nm shortpass (E700SP-2, Chroma) and a 615 nm bandpass filter (ET615/30m-2P, Chroma) and detected by a HPM-100-40 detector (Becker & Hickl) containing a hybrid Hamamatsu R10467-40 GaAsP photomultiplier tube. Images of 256 × 256 pixels (256 time bins/pixel) were obtained and averaged over 10 acquisitions of 60 s. Photon data were analyzed using a combination of custom MATLAB (The Mathworks) scripts and SPCImage software (Becker & Hickl).

Synthesis

Dibenzoazacyclooctyne-647. A flame-dried RB flask containing a stir bar was charged with dibenzoazacyclooctyne amine (0.440 mg, 1.59 μmol), and dry DMF (44.0 μL). A solution of

Alexa Fluor 647 carboxylic acid, succinimidyl ester (1.00 mg) in dry DMF (456 μL) was added dropwise. To this suspension, DIPEA (0.700 μL, 4.00 μmol) was then added, and the mixture was stirred overnight at rt. After concentration, the crude solid was purified by reversed-phase HPLC using a gradient of 0% to 40% MeCN in H_2O (0.1% TFA) over 45 min. The collected fractions were concentrated and lyophilized to yield a blue solid (1.10 mg). See ^{1}H NMR spectrum below. HRMS (ESI): $[M]^{3-}$ found, *m/z* = 308.7328. λ_{max} = 648 nm.

Fab-594. Anti-integrin $\alpha_V\beta_3$ IgG (200 μg, LM609 clone) was concentrated and buffer exchanged into digestion buffer (Pierce), pH 10.0 with a 30k MWCO centrifugal filter (EMD Millipore) and treated with immobilized papain (Pierce) for 5.5 h at 37 ^oC. After filtration, unreacted IgG and Fc fragments were removed from the reaction mixture by treatment with immobilized protein A (Thermo) for 30 min at rt, followed by filtration. Cleavage and purification were assessed by SDS-PAGE under reducing conditions and indicated the presence of only Fab. Purified Fab was concentrated and buffer exchanged into PBS with a 10k MWCO centrifugal filter (EMD Millipore). After addition of Alexa Fluor 594 carboxylic acid, succinimidyl ester (3.67 μg, 4.48 nmol) and 100 mM sodium bicarbonate, the mixture was rotated overnight at rt. Fab-594 was purified and concentrated with a 10k MWCO centrifugal filter, yielding a light pink solution (22.1 μM, DOL = 1.06). λ_{max} = 590 nm.

Cell culture

U87MG cells were cultured in DMEM media without phenol red and supplemented with 10% FBS, 100 units/mL penicillin and 100 μg/mL streptomycin. The cells were maintained at 37 °C and 5% CO_2 in a water-saturated incubator and counted using a hemocytometer.

U87MG lysate

U87MG cells were grown to 80-90% confluency on a 145 mm tissue culture dish. The cells were washed with two portions of PBS (10 mL). Lysis buffer (20 mM Tris-Cl, pH 7.5, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, Roche protease inhibitors) was added and the cells were mechanically agitated. Lysate was then incubated at 4 °C for 20 min, followed by centrifugation at 12,000 x g for 40 min. The supernatant was removed, and the pellet was discarded. Protein concentration was determined using the BCA protein assay (Pierce), and the protein content was normalized across samples.

Lectin blot

U87MG lysate was first treated with Protein A/G Ultralink Resin (Pierce) for 1 h at 4 °C. After centrifugation (2,500 x g, 3 min), the supernatant was removed and rotated overnight at 37 °C with or without neuraminidase (1000 units, *Clostridium perfringens*, New England Biolabs). Anti-integrin $\alpha_V\beta_3$ antibody (5 μg) was subsequently added. The mixture was then rotated overnight at 4 °C. To this mixture, Protein A/G Ultralink Resin was administered and allowed to rotate for 4 h at 4 °C. Centrifugation was performed and the resin was washed (3X) with lysis buffer. Loading buffer was added to the resin and boiled at 95 °C for 5 min. SDS-PAGE was performed, followed by transfer to a nitrocellulose membrane. The membrane was first stained for protein with Ponceau. The membrane was then blocked with blocking buffer (0.5% Tween-

20, DPBS containing calcium and magnesium) for 1 h at rt. The presence of sialic acid was probed with either SNA- or MAA-fluorescein lectin conjugates by incubating the membrane with either lectin at 1:1000 dilution in blocking buffer for 1 h at rt. After washing the membrane with blocking buffer (3X), the blot was imaged for fluorescence.

Integrin $\alpha_V\beta_3$ azide incorporation

U87MG cells were cultured with either Ac_4ManNAz or Ac_4ManNAc (100 μM) for 3 days in a 145 mm petri dish. After washing the cells twice with PBS (10 mL), lysates were formed according to the above procedure. Lysates were first treated with Protein A/G Ultralink Resin (Pierce) for 1 h at 4 °C. After centrifugation (2,500 x g, 3 min), the supernatant was removed and rotated overnight at 37 °C with or without neuraminidase (1000 units, *Clostridium perfringens*, New England Biolabs), followed by treatment with phosphine-biotin (250 μM) overnight at 37 °C. Anti-integrin $\alpha_V\beta_3$ antibody (5 μg) was added. The mixture was then rotated overnight at 4 °C. To this mixture, Protein A/G Ultralink Resin was administered and allowed to rotate for 4 h at 4 °C. Centrifugation was performed and the resin was washed (3X) with lysis buffer. Loading buffer was added to the resin and boiled at 95 °C for 5 min. SDS-PAGE was performed, followed by transfer to a nitrocellulose membrane. The membrane was first stained for protein with Ponceau. The membrane was then blocked with blocking buffer (5% BSA in 0.05% Tween-20, DPBS) for 2 h at rt. The presence of biotin was probed with an anti-biotin, Alexa Fluor 647 antibody (1:3400 dilution, Jackson ImmunoResearch Laboratories) in blocking buffer for 1 h at rt. After washing the membrane with 0.05% Tween-20, DPBS (3X), the blot was imaged for fluorescence.

Fluorescence microscopy

U87MG cells were cultured on slides mounted with 8-tissue culture wells in 0.3 mL containing Ac_4ManNAz or Ac_4ManNAc (100 μM). After 3 d, the cells were washed with DPBS + 1% FBS (2 x 0.5 mL). To examine the labeling of SiaNAz residues, cells were then treated with DIBAC-647 (25 μM) in media for 1 h at 4 °C. The cells were then washed with DPBS + 1% FBS (3 x 0.4 mL) and fixed with 4% paraformaldehyde in PBS (0.2 mL) for 5 min at 4 °C and 15 min at rt. After washing with DPBS (4 x 0.2 mL), the cells were imaged in DPBS (0.2 mL). To investigate Fab-594 specificity, cells were fixed with 4% paraformaldehyde in PBS (0.2 mL) for 5 min at 4 °C and 15 min at rt; the cells were then washed with DPBS + 1% FBS (4 x 0.2 mL). The cells were treated with either Fab-594 (500 nM) or Fab-594 and anti- $\alpha_V\beta_3$ IgG (500 nM each) in DPBS for 1 h at 4 °C. After washing with DPBS (3 x 0.4 mL), the cells were imaged in DPBS (0.2 mL).

Fluorescence lifetime imaging microscopy (FLIM)

In a typical experiment, U87MG cells were cultured on slides mounted with 8-tissue culture wells in 0.3 mL containing Ac_4ManNAz or Ac_4ManNAc (100 μM). After 3 d, the cells were washed with DPBS + 1% FBS (2 x 0.5 mL). The cells were then treated with DIBAC-647 in media for 1 h at 4 °C. After washing the cells with DPBS + 1% FBS (3 x 0.4 mL), the cells were fixed with 4% paraformaldehyde in PBS (0.2 mL) for 5 min at 4 °C and 15 min at rt. After

washing with DPBS (4 x 0.2 mL), the cells were incubated with Fab-594 (500 nM), followed by washing with DPBS (3 x 0.4 mL). The cells were imaged at rt in DPBS (0.2 mL).

Analysis of the acquired FLIM data was performed by first binning (3 pixels x 3 pixels) the time-dependent photon image and assigning a minimum threshold count of 50 recorded photons for modeling. Exponential fits (one-component) are obtained for the decay histogram with limits of 1 ns and 8 ns. Fluorescence lifetime values derived from the exponential fits are then displayed as a heat-map image. Cell surface lifetime values are also extracted using a search algorithm and displayed as a histogram.

Metabolic labeling of human tissue slice cultures

Informed, signed consent was obtained from either the patient or next of kin for human samples. An 8 mm diameter core from a radical prostatectomy was embedded in agarose and segmented into thin slices (300 um thick) with an automated Krumdieck microtome. Tissue sections were retrieved, frozen, and evaluated for histopathology grading using hematoxylin and eosin stains. These slices were loaded onto specially designed titanium screens and transferred to standard 6-well culture plates containing 2.5 ml of serum-free medium (PFMR-4A) supplemented with 10 ng/mL cholera toxin, 10 ng/mL epidermal growth factor, 40 μg/mL bovine pituitary extract, 100 μM phosphoethanolamine, 3 μM hydrocortisone, 30 nM selenous acid, 100 μg/mL gentamycin, 0.03 nM all-trans retinoic acid, 4 μg/mL insulin, 2.3 μM α-tocopherol, and 50 nM synthetic androgen R1881 (in EtOH) and 50 μM Ac_4ManNAz or Ac_4ManNAc. The dishes were rotated (1 rpm) on an inclined plane (30°) in a tissue culture incubator to permit full oxygenation of the tissue. Each day, the conditioned media was removed and replaced with fresh media. After three days, the conditioned media was again removed, and the slices washed with cold PBS prior to being snap-frozen in liquid nitrogen for storage.

Human prostate tissue slice lysate

Prostate tissue slices were lysed in 1 mL of lysis buffer containing 20 mM Tris-Cl, pH 7.4, 300 mM NaCl, 1% Triton X-100, and protease inhibitors (Calbiochem Inhibitor Cocktail III, EMD Millipore). The tissue was pulverized at 4 °C using a Tissue Tearor (setting 2, Biospec Products, Model #780CL-04) with cycles of 20 sec on/20 sec off for 3 min. Lysate was incubated at 4 °C for 30 min and the supernatant was collected after centrifugation at 10,000 rpm for 10 min at 4 ^{o}C. The supernatant was collected, and a BCA protein assay (Pierce) was performed to determine protein concentration. The protein content was normalized across samples.

Prostate tissue lectin blot

Prostate tissue lysate was first treated with Protein A/G Ultralink Resin (Pierce) for 1 h at 4 °C. After centrifugation (2,500 x g, 3 min), the supernatant was removed and anti-integrin $\alpha_V\beta_3$ antibody (5 μg) was subsequently added. The mixture was then rotated overnight at 4 °C. To this mixture, Protein A/G Ultralink Resin was administered and allowed to rotate for 4 h at 4 °C. Centrifugation was performed and the resin was washed (3X) with lysis buffer. Loading buffer was added to the resin and boiled at 95 °C for 5 min. SDS-PAGE was performed, followed by transfer to a nitrocellulose membrane. The membrane was first stained for protein with Ponceau. The membrane was then blocked with blocking buffer (0.5% Tween-20, DPBS containing

calcium and magnesium) for 1 h at rt. The presence of sialic acid was probed with a MAA-fluorescein lectin conjugate by incubating the membrane with the lectin at 1:1000 dilution in blocking buffer for 1 h at rt. After washing the membrane with blocking buffer (3X), the blot was imaged for fluorescence.

Prostate tissue azide incorporation

Prostate tissue lysates were formed according to the above procedure. The lysates were first treated with Protein A/G Ultralink Resin (Pierce) for 1 h at 4 °C. After centrifugation (2,500 x g, 3 min), the supernatant was removed and rotated overnight at 37 °C with or without neuraminidase (1000 units, *Clostridium perfringens*, New England Biolabs), followed by treatment with phosphine-biotin (250 μM) overnight at 37 °C. Anti-integrin $\alpha_V\beta_3$ antibody (5 μg) was added. The mixture was then rotated overnight at 4 °C. To this mixture, Protein A/G Ultralink Resin was administered and allowed to rotate for 4 h at 4 °C. Centrifugation was performed and the resin was washed (3X) with lysis buffer. Loading buffer was added to the resin and boiled at 95 °C for 5 min. SDS-PAGE was performed, followed by transfer to a nitrocellulose membrane. The membrane was first stained for protein with Ponceau. The membrane was then blocked with blocking buffer (5% BSA in 0.05% Tween-20, DPBS) for 2 h at rt. The presence of biotin was probed with an anti-biotin, Alexa Fluor 647 antibody (1:3400 dilution, Jackson ImmunoResearch Laboratories) in blocking buffer for 1 h at rt. After washing the membrane with 0.05% Tween-20, DPBS (3X), the blot was imaged for fluorescence.

Prostate adenocarcinoma tissue slice fluorescence microscopy

Frozen tissue slices were placed on slides mounted with 4-tissue culture wells and allowed to thaw. Tissue slices were then washed with DPBS + 1% FBS (2 x 0.8 mL), followed by incubation with or without DIBAC-647 (25 μM) in DMEM media for 1 h at 4 °C. Tissues slices were subsequently washed with DPBS + 1% FBS (10 x 0.8 mL) and fixed with 4% paraformaldehyde in PBS (0.4 mL) for 5 min at 4 °C and 20 min at rt. After washing the tissues slices with DPBS (4 x 0.4 mL), the tissue slices were blocked with 10% Normal Goat Serum in DPBS (0.8 mL) for 1 h at rt. The blocking solution was removed, and the tissue slices were then treated with either Fab-594 (500 nM), Fab-594 and anti-$\alpha_V\beta_3$ IgG (500 nM each), or DPBS for 1 h at 4 °C. The tissue slices were washed with DPBS (3 x 0.8 mL) and transferred to a concave microscope slide. After addition of DPBS (50 uL), a glass coverslip was used to seal the slide, and the tissue slices were imaged.

Prostate adenocarinoma tissue slice FLIM

In a typical experiment, frozen tissue slices were placed on slides mounted with 4-tissue culture wells and allowed to thaw. Tissue slices were then washed with DPBS + 1% FBS (2 x 0.8 mL), followed by incubation with DIBAC-647 (25 μM) in DMEM media for 1 h at 4 °C. Tissues slices were subsequently washed with DPBS + 1% FBS (10 x 0.8 mL) and fixed with 4% paraformaldehyde in PBS (0.4 mL) for 5 min at 4 °C and 20 min at rt. After washing the tissues slices with DPBS (4 x 0.4 mL), the tissue slices were blocked with 10% Normal Goat Serum in DPBS (0.8 mL) for 1 h at rt. The blocking solution was removed, and the tissue slices were then treated with Fab-594 (500 nM) in DPBS for 1 h at 4 °C. The tissue slices were washed

with DPBS (3 x 0.8 mL) and transferred to a concave microscope slide. After addition of DPBS (50 uL), a glass coverslip was used to seal the slide.

References

(1) Liu, Y.-C.; Yen, H.-Y.; Chen, C.-Y.; Chen, C.-H.; Cheng, P.-F.; Juan, Y.-H.; Chen, C.-H.; Khoo, K.-H.; Yu, C.-J.; Yang, P.-C.; Hsu, T.-L.; Wong, C.-H. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11332 - 11337.

(2) Partridge, E. A.; Le Roy, C.; Di Guglielmo, G. M.; Pawling, J.; Cheung, P.; Granovsky, M.; Nabi, I. R.; Wrana, J. L.; Dennis, J. W. *Science* **2004**, *306*, 120 - 124.

(3) Ohtsubo, K.; Takamatsu, S.; Minowa, M. T.; Yoshida, A.; Takeuchi, M.; Marth, J. D. *Cell* **2005**, *123*, 1307 - 1321.

(4) Cha, S.-K.; Ortega, B.; Kurosu, H.; Rosenblatt, K. P.; Kuro-o, M.; Huang, C.- L. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 9805 - 9810.

(5) Haltiwanger, R. S.; Lowe, J. B. *Annu. Rev. Biochem.* **2004**, *73*, 491 - 537.

(6) Marth, J. D.; Grewal, P. K. *Nat. Rev. Immunol.* **2008**, *8*, 874 - 887.

(7) Fuster, M. M.; Esko, J. D. *Nat. Rev. Cancer* **2005**, *5*, 526 - 542.

(8) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125 - 1128.

(9) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007 - 2010.

(10) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664 - 667.

(11) Conze, T.; Carvalho, A. S.; Landegren, U.; Almeida, R.; Reis, C. A.; David, L.; Söderberg, O. *Glycobiology* **2010**, *20*, 199 - 206.

(12) Haga, Y.; Ishii, K.; Hibino, K.; Sako, Y.; Ito, Y.; Taniguchi, N.; Suzuki, T. *Nat. Commun.* **2012**, *3*, 907.

(13) Galbraith, C. G.; Galbraith, J. A. *J. Cell Sci.* **2011**, *124*, 1607 - 1611.

(14) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. *Essentials of Glycobiology*, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2008.

(15) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046 - 15047.

(16) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16793 - 16797.

(17) Levitt, J. A.; Matthews, D. R.; Ameer-Beg, S. M.; Suhling, K. *Curr. Opin. Biotechnol.* **2009**, *20*, 28 - 36.

(18) Meyer, A.; Auernheimer, J.; Modlinger, A.; Kessler, H. *Curr. Pharm. Des.* **2006**, *12*, 2723 - 2747.

(19) Kremser, M. E.; Przybyło, M.; Hoja-Łukowicz, D.; Pocheć, E.; Amoresano, A.; Carpentieri, A.; Bubka, M.; Lityńska, A. *Biochim. Biophys. Acta Gen. Subj.* **2008**, *1780*, 1421 - 1431.

(20) Chiodelli, P.; Urbinati, C.; Mitola, S.; Tanghetti, E.; Rusnati, M. *J. Biol. Chem.* **2012**, *287*, 20456 - 20466.

(21) Markowska, A. I.; Liu, F.-T; Panjwani, N. *J. Exp. Med.* **2010**, *207*, 1981 - 1993.

(22) Zhang, X.; Xiong, Z.; Wu, Y.; Cai, W.; Tseng, J. R.; Gambhir, S. S.; Chen, J. X. *Nucl. Med.* **2006**, *47*, 113 - 121.

(23) Benson, R. C.; Meyer, R. A.; Zaruba, M. E.; McKhann, G. M. *J. Histochem. Cytochem.* **1979**, *27*, 44 - 48.

(24) Belardi, B.; O'Donoghue, G. P.; Smith, A.W.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9549 - 9552.

(25) Hubbard, S. C.; Boyce, M.; McVaugh, C. T.; Peehl, D. M.; Bertozzi, C. R. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4945 - 4950.

(26) Goel, H. L.; Li, J.; Kogan, S.; Languino, L. R. *Endocr.-Relat. Cancer* **2008**, *15*, 657 - 664.

(27) Laughlin, S. T.; Bertozzi, C. R. *Nat. Protocols* **2007**, *2*, 2930 - 2944.

Chapter 5

Glycomic Signatures Regulate Nuclear Galectin-1 to Pattern the Mammary Gland

Chapter 5. Glycomic Signatures Regulate Nuclear Galectin-1 to Pattern the Mammary Gland[a]

Abstract

Branching morphogenesis in the mammary gland is achieved by the invasion of epithelial cells through a microenvironment consisting of stromal cells and dense extracellular matrix (ECM)[1,2]. Here we demonstrate that Galectin-1 (Gal-1), an endogenous lectin that recognizes *N*-acetyllactosamine (LacNAc)-containing glycoconjugates, induces invasion and branching of mammary epithelia *in vivo* and in 3D organotypic cultures. The positive regulation of branching and invasion by Gal-1 was independent of its glycan-binding ability and instead was dependent on its localization within mammary epithelial nuclei. The nuclear translocation of Gal-1 was however dependent on a glycomic signature that was unique to the front of the mammary end bud. α2,6-sialylation (α2,6-SA) of terminal LacNAc residues in the end buds masked Gal-1 ligands and promoted nuclear localization of Gal-1 and thereby epithelial invasiveness. Conversely, terminal LacNAc glycans, enriched in the epithelial ducts, sequestered Gal-1 in the extracellular environment, ultimately attenuating invasive potential. These findings suggest that the two glycomic signatures at the level of tissue microanatomy regulate the nuclear function of Gal-1. Our results also provide a novel glycan-lectin based mechanism through which the tissue microenvironment modulates the invasive behavior of mammary epithelia in a developmental context, suggesting a possible explanation of how malignant breast cells, which are rich in α2,6-SA[3] and Gal-1,[4] could invade.

Introduction

Gal-1 belongs to a small set of proteins that lacks a signal peptide but is secreted to the extracellular environment through unconventional mechanisms[5]. There, it interacts with glycoconjugates, modulating their surface organization and mediating cell-cell contact[6]. Within the cell, Gal-1 localizes to both the cytosol and the nucleus, where it appears to function independent of its glycan recognition[7]. To date, we have no explanation for the presumed distinct functions of Gal-1 in different subcellular contexts.

We and others have shown that tissue architecture is a dominant regulator of proliferation and invasion of cancer cells through their surrounding microenvironment[8-10]. While glycomic changes such as hypersialylation have long been shown to correlate with cell invasion, how glycan patterns and lectins mechanistically drive the invasion process remains obscure. Mammary epithelia also proliferate and invade their surrounding stroma during mammary gland development, although in a more controlled manner. Because Gal-1 is upregulated in invasive breast cancer, and tumorigenesis involves a radical change in glandular architecture, we hypothesized that Gal-1 could play a crucial role during the development of the mammary gland.

[a] Ramray Bhat contributed to work presented in this chapter.

Results and Discussion

Gal-1 is concentrated in mammary gland end buds

Using immunofluorescence, we measured the levels of Gal-1 in murine mammary glands at distinct stages of development. Gal-1 levels were found to be highest during the early stages of branching morphogenesis (5 weeks post partum) (Figure 5.1a; Figure 5.2 shows weak to moderate levels of Gal-1 at other developmental stages of the mammary gland). Within 5-week glands, Gal-1 expression was highest in epithelial cells at the front of the end buds: bulbous microanatomical features that represent the invading fronts of the mammary arbor during the branching stage of development (Figure 5.1b). Gal-1 levels were low in the quiescent, noninvasive epithelia of mammary ducts even at this stage (Figure 5.1b). We also found a major difference in subcellular localization: Gal-1 in the end bud epithelia was predominantly nuclear, whereas, in the ductal cells, it was depleted from nuclei and instead showed weak levels in the cytosol and glycocalyx (Figure 5.1b). To understand whether Gal-1 levels were linked to the ability of mammary epithelia to invade into, and branch within, the mammary fat pad, we used laser scanning-based tissue autofluorescence/fluorescence imaging (LS-TAFI), an imaging technique that measures autofluorescence of thick tissues and accurately reconstructs three dimensional microanatomy[11]. Mammary gland whole-mounts from 5-week Gal-1$^{-/-}$ mice showed a decrease in branching and invasion of mammary epithelia within their fat pads compared to wildtype mammary glands (Figure 5.1c).

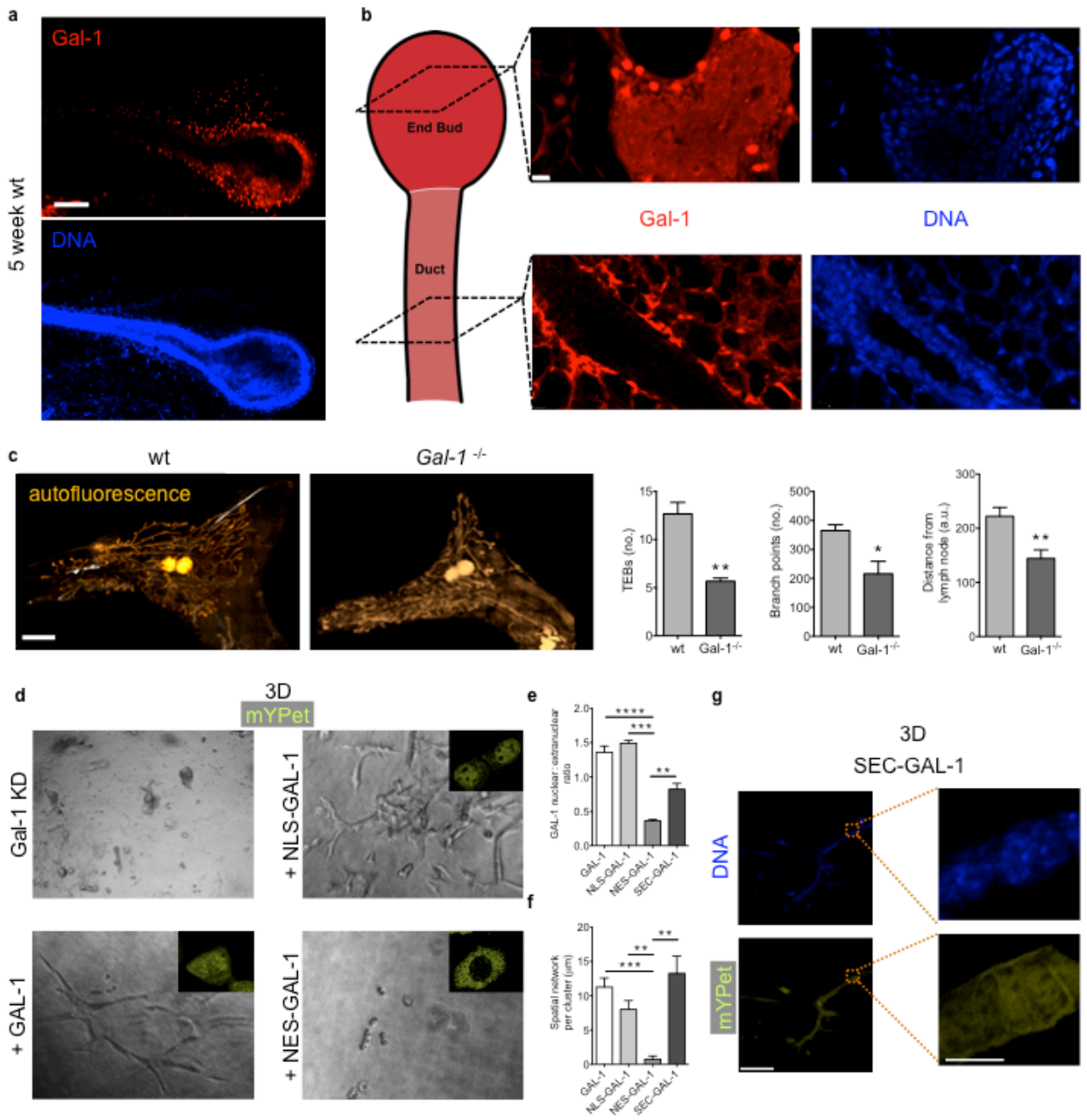


Figure 5.1. Nuclear Gal-1 drives mammary invasion and branching. **a**, Immunofluorescence micrographs of a five-week murine mammary gland stained with α-Gal-1 antibody (top) and DAPI (bottom). Scale bar, 100 µm. **b**, In the end bud, Gal-1 is enriched in the nucleus (top), whereas Gal-1 staining in the mammary duct is mainly extracellular (bottom). Scale bar, 10 µm. **c**, Autofluorescence micrographs of wildtype (left) and Gal-1$^{-/-}$ (right) murine mammary gland showing epithelial architecture (left panels). Scale bar, 1500 µm. Quantification of the number of terminal end buds (left graph), the number of branch points (middle graph), and the distance from lymph node (right graph) per gland for wildtype and Gal-1$^{-/-}$ murine mammary networks (N=3). **d**, Gal-1 KD Eph4 cells (top, left) ectopically expressing either NLS-GAL-1 (top, right), GAL-1 (bottom, left), or NES-GAL-1 (bottom, right) were cultured in a 3D collagen I (CL-I)

gel. Branching was observed only upon expression of GAL-1 or NLS-GAL-1. (Images taken with 20x objective). Inset fluorescence micrographs indicate the subcellular localization of each mYPet fusion construct. **e**, Quantification of GAL-1 nuclear:extranuclear ratio for each of the GAL-1 constructs. **f**, Quantification of the spatial network per cluster for each of the GAL-1 constructs in 3D. Spatial network is defined as the sum of the branch lengths for each cluster. **g**, Gal-1 KD EpH4 cells ectopically expressing SEC-GAL-1 branch when cultured in a 3D CL-I gel (left panel). Scale bar, 100 µm. mYPet fluorescence of SEC-GAL-1 fusion construct is distributed between the extracellular space and the nucleus (right panel). Scale bar, 10 µm. For all bar graphs, error bars represent s.e.m. Statistical significance is given by $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$.

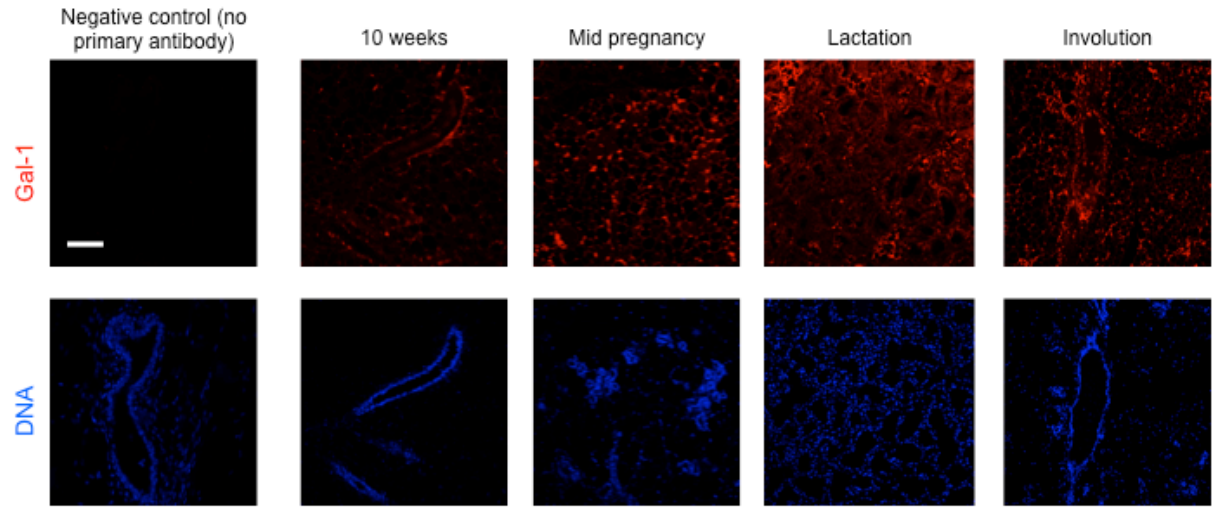


Figure 5.2 Gal-1 levels at different developmental stages of mammary gland morphogenesis. Immunofluorescence micrographs of paraffin embedded sections of murine mammary glands during distinct stages of development, at 5 weeks but omitting the primary antibody, at 10 weeks, midpregnancy, lactation, and involution stained with α-Gal-1 antibody (upper panel) and DAPI (lower panel). Gal-1 levels are low after branching is complete and during pregnancy but increase moderately during lactation and during involution. Scale bar, 100 µm.

Nuclear Gal-1 drives mammary gland branching and invasion

In order to probe whether the nuclear localization of Gal-1 at the end bud was required for epithelial invasion and branching, we relied on an organotypic 3D culture model of mammary epithelial cells. Upon addition of epidermal growth factor[12-15], EpH4 mammary epithelial cells, embedded in collagen Type-1 (CL-1) matrix, invade into their surroundings and form branched structures. Having confirmed the presence of endogenous Gal-1 within invading epithelia in 3D culture (Figure 5.3), we knocked down (KD) Gal-1 (Figure 5.4) and observed an abrogation of branching and invasion (Figure 5.1d). We were able to rescue the invasive phenotype by overexpression of human Gal-1 (GAL-1). Having observed nuclear enrichment of Gal-1 in the end buds *in vivo*, we sought to evaluate invasion/branching in culture by overexpressing GAL-1 constructs tagged with specific localization signals: a nuclear localization signal (NLS) and a nuclear export signal (NES). The NLS tag directed nuclear localization of

GAL-1 and was able to rescue branching and invasion (Figure 5.1d and Figure 5.5), while expression of NES-GAL-1 resulted in the localization of GAL-1 to the cytoplasm and failed to rescue branching (Figure 5.1d and Figure 5.5).

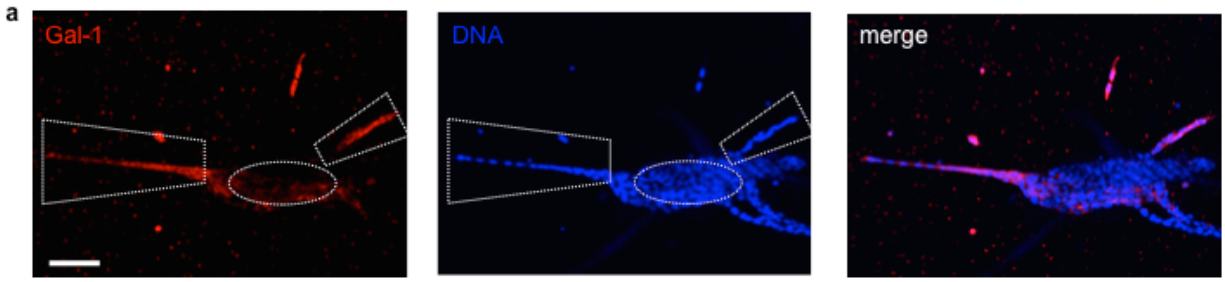


Figure 5.3. Gal-1 is concentrated in invading branches of EpH4 cells in 3D culture. Wildtype EpH4 cells exhibit branching and invasion when cultured in 3D CL-I gels and stained with α-Gal-1 antibody and DAPI. The epithelia constituting the branches (bounded by trapezoid) show particularly high levels of Gal-1 when compared with the non-branching core (bounded by oval). Scale bar, 70 μm.

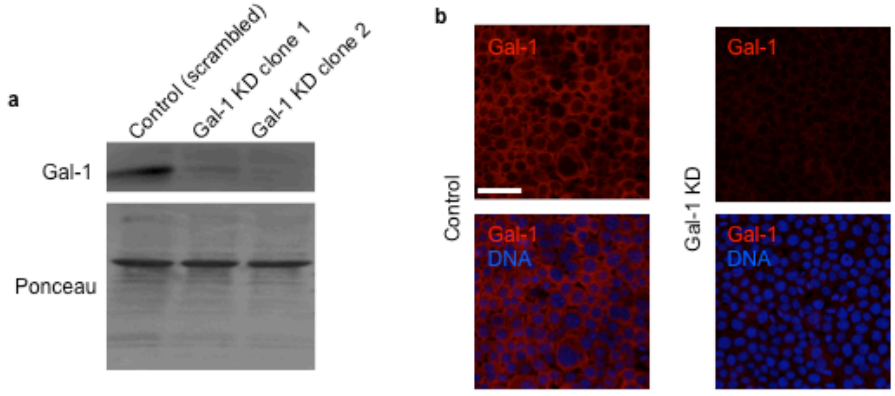


Figure 5.4. shRNA-based Gal-1 knockdown. **a**, Immunoblot showing Gal-1 knockdown in lysates from EpH4 cells using one scrambled control shRNA and two separate shRNA clones against Gal-1. Ponceau staining of membrane indicates uniform loading. **b**, Immunofluorescence micrographs of control cells and cells transduced with Gal-1 shRNA lentivirus showing knockdown of Gal-1. Scale bar, 50 μm.

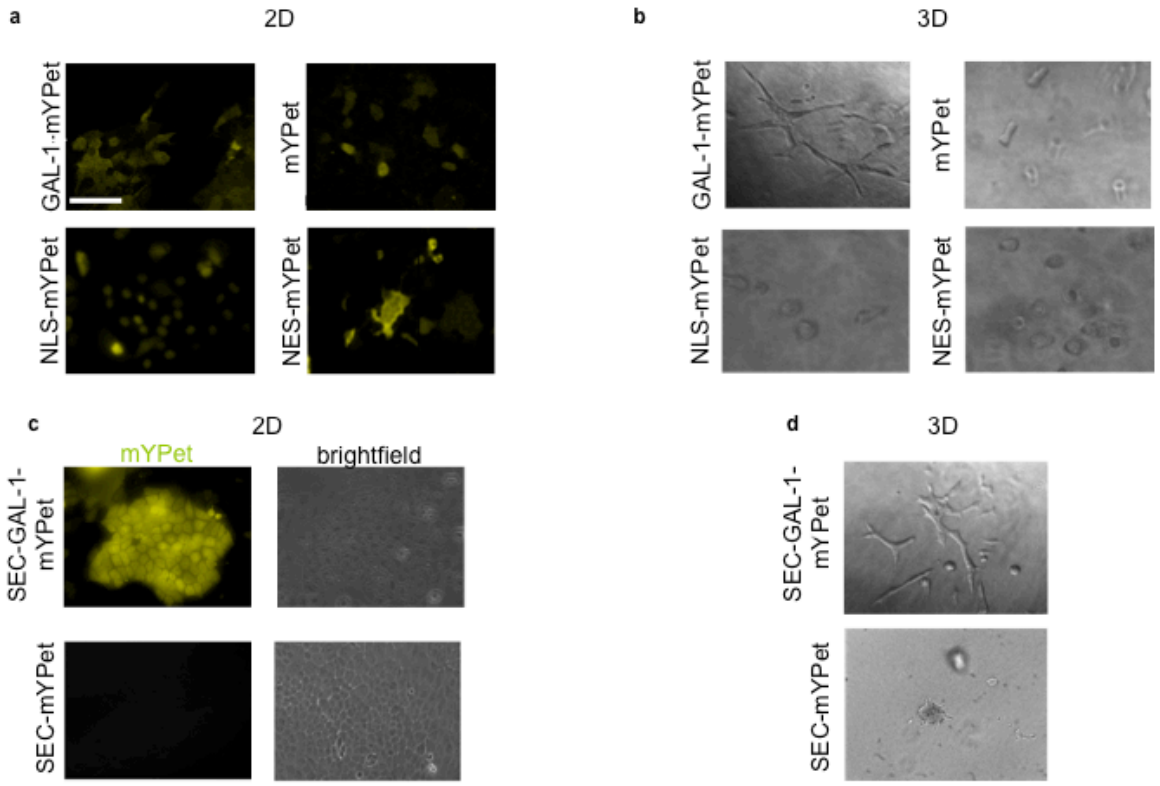


Figure 5.5. Localization of subcellular constructs and branching of EpH4 cells expressing subcellular constructs. **a-b,** Gal-1 KD Eph4 cells (top, left) ectopically expressing GAL-1 (top, right), NLS (bottom, left), or NES (bottom, right) tagged with mYPet or only mYPet (top, right) were cultured in 2D and in (**b**) 3D . **c-d**, Gal-1 KD Eph4 cells ectopically expressing SEC-GAL-1 (top) fused to mYpet and SEC-mYpet (bottom) were cultured in (**c**) 2D and in (**d**)3D. Scale bar, 50 μm.

Endogenous Gal-1 is known to be secreted to the extracellular space through an unconventional secretion pathway[16]. But, we wanted to ensure that once translated within mammary epithelia, Gal-1 is first secreted to the extracellular space before it has the opportunity to traffic elsewhere. Therefore, we overexpressed GAL-1 with an extracellular secretion signal peptide (SEC) tag. SEC-GAL-1 was secreted and then relocalized back within the nucleus, and also rescued branching (Figure 5.1e; Figure 5.5, Figure 5.6 for confirmation of the nuclear localization of SEC-GAL-1 within mammary epithelia cocultured with SEC-GAL-1 expressing cells, and Figure 5.7 showing nuclear localization of exogenously added GAL-1). These data together suggest that nuclear Gal-1 is necessary and sufficient for invasion and branching and that Gal-1 is able to translocate from the extracellular space to the nucleus.

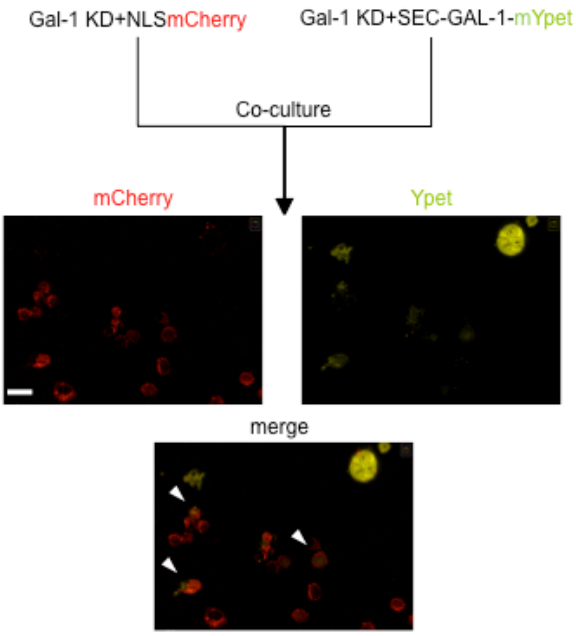


Figure 5.6. SEC-GAL-1 translocates to mammary epithelial nucleus in adjacent cells. Gal-1 KD EpH4 cells ectopically expressing NLS-mCherry were co-cultured with GAL-1 KD EpH4 cells ectopically expressing SEC-GAL-1 fused to mYpet. Fluorescent micrographs show colocalization of both mCherry and mYpet (white arrowheads). Scale bar, 50 µm.

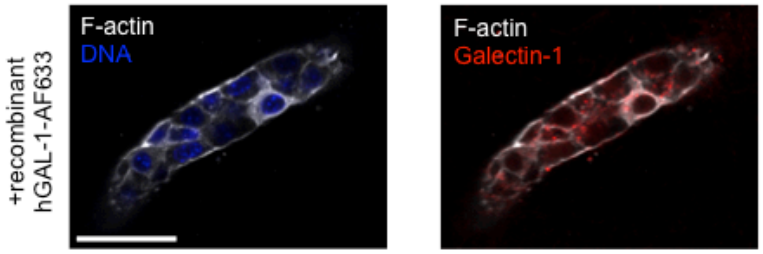


Figure 5.7. Nuclear translocation of exogenous GAL-1. Immunofluorescence micrographs of Gal-1 KD Eph4 cells cultured in 3D CL-I gels with exogenous GAL-1-Alexa Fluor 633 showing branching when stained with F-actin and DAPI. Scale bar, 50 µm.

Galectin-glycan recognition has previously been reported to influence cell invasion and migration[17]. Consequently, we investigated if mammary epithelial invasion was at all dependent on Gal-1's interaction with its cognate glycan ligands by constructing the N46D mutant of GAL-1, which attenuates glycan binding[18]. Overexpression of GAL-1 (N46D) in Gal-1 KD EpH4 cells rescued the branching phenotype in 3D similar to wildtype Gal-1 (Figure 5.8a). When these same cells were cultured in 2D, the mutant GAL-1 showed a greater degree of nuclear localization than wildtype Gal-1 (Figure 5.8b,c). This led us to hypothesize that cognate glycan ligands of Gal-1 may modulate the subcellular distribution of Gal-1 but are not required for Gal-1's effect on branching. As well, we noted that endogenous nuclear Gal-1 levels were higher in cells

cultured in 3D, compared to cells cultured in 2D, emphasizing the importance of the cellular microenvironment in regulating the subcellular localization of Gal-1 (Figure 5.8d,e).

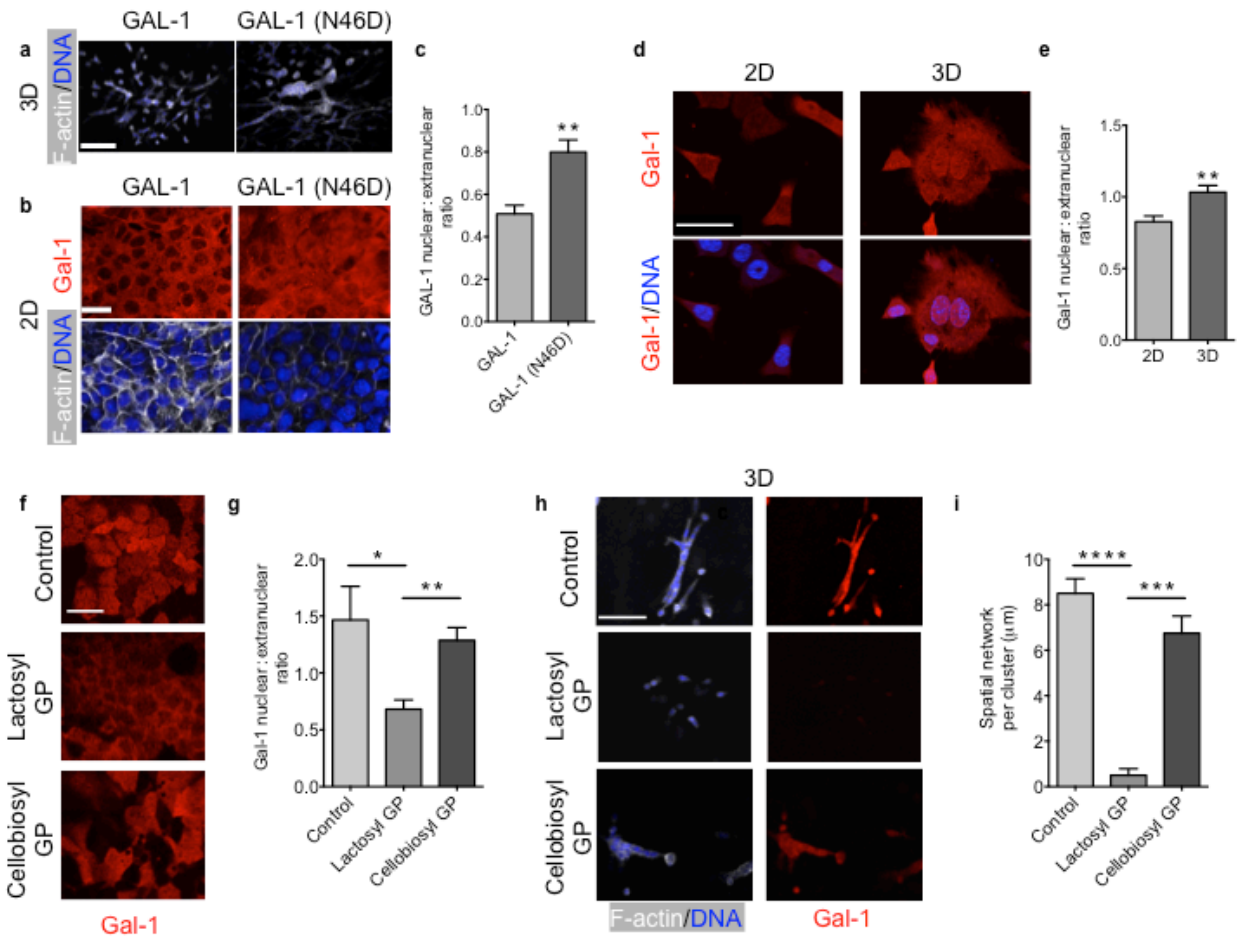


Figure 5.8. Nuclear localization of Gal-1 is dependent on extracellular glycan ligands. **a-b**, Fluorescence micrographs of Gal-1 KD EpH4 cells ectopically expressing GAL-1 (left panel) and GAL-1 (N46D) (right panel) fusion proteins in 3D (**a**) or 2D (**b**) and stained with F-actin and DAPI. Cells expressing GAL-1 (N46D), a mutant with attenuated glycan binding, invade and branch when cultured in 3D. Scale bar, 100 μm. **c**, Quantification of GAL-1 nuclear:extranuclear ratio shows GAL-1 (N46D) is concentrated in the nucleus. **d**, Immunofluorescence micrographs of EpH4 cells cultured in 2D (left panel) or in a 3D CL-I gel (right panel) and stained with an α-Gal-1 antibody and DAPI. Scale bar, 25 μm. **e**, Quantification of Gal-1 nuclear:extranuclear ratio for EpH4 cells cultured in 2D vs. 3D. **f**, Fluorescence micrographs of GAL-1 expressing EpH4 cells grown on top of CL-I gels treated with either a soluble lactosyl GP, which binds Gal-1, or a soluble cellobiosyl GP, which does not interact with Gal-1. Scale bar, 50 μm. **g**, Quantification of the GAL-1 nuclear:extranuclear ratio for EpH4 cells treated with either the lactosyl GP or the cellobiosyl GP. **h**, Immunofluoresence micrographs of EpH4 cells cultured in 3D CL I gel in the presence or absence of either lactosyl GP or cellobiosyl GP and then washed and stained with F-actin, DAPI and α-Gal-1 antibody. Scale bar, 100 μm. **i**, Quantification of the spatial network per cluster of EpH4 cells in the presence or absence of GPs. For all bar graphs, error bars represent s.e.m. Statistical significance is given by $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$.

Gal-1 subcellular localization is dependent on glycan microenvironment

What controls the relative abundances of Gal-1 in the different cellular compartments? To answer this, we re-engineered the glycan microenvironment surrounding the cells using glycopolymers that mimic ECM glycoproteins[19]. We generated glycopolymers (GPs) functionalized with Gal-1 ligands (lactosyl GP) or control glycan structures (cellobiosyl GP) that do not associate with Gal-1. When lactosyl GP was incubated with mammary epithelial cells cultured on top of collagen, we found a marked decrease in nuclear GAL-1 after 1 day (Figure 5.8f,g). In contrast, untreated cells and cells treated with cellobiosyl GP showed greater levels of nuclear GAL-1. When lactosyl GP was added to 3D CL-I cultures, we found an abrogation of branching/invasive phenotype (Figure 5.8h,i). Gal-1 was not visible after washing lactosyl GP-treated cells. In contrast, Gal-1 staining was evident in branching cultures of cells treated with cellobiosyl GP.

The addition of α2,6-SA residues is known to negatively regulate Gal-1's binding to LacNAc epitopes[20,21]. We found that engineering the levels of extracellular sialic acid residues capping terminal LacNAc structures increases the translocation of Gal-1 to the nucleus. Through the use of two independent methods, exogenous addition of Ac_4ManNAc, a metabolic precursor of sialic acid, and overexpression of GNE, the rate-limiting enzyme in sialic acid synthesis, we were able to vary the extent of cell surface sialylation[22] (Figure 5.9a). In both methods, α2,6-SA levels, measured by staining of *Sambucus nigra* lectin (SNA)[22] increased (Figure 5,9b) and so did the nuclear Gal-1 levels (Figure 5.9c). We also knocked down GNE with shRNA in EpH4 cells, which caused a decrease in branching and invasion in 3D (Figure 5.9d).

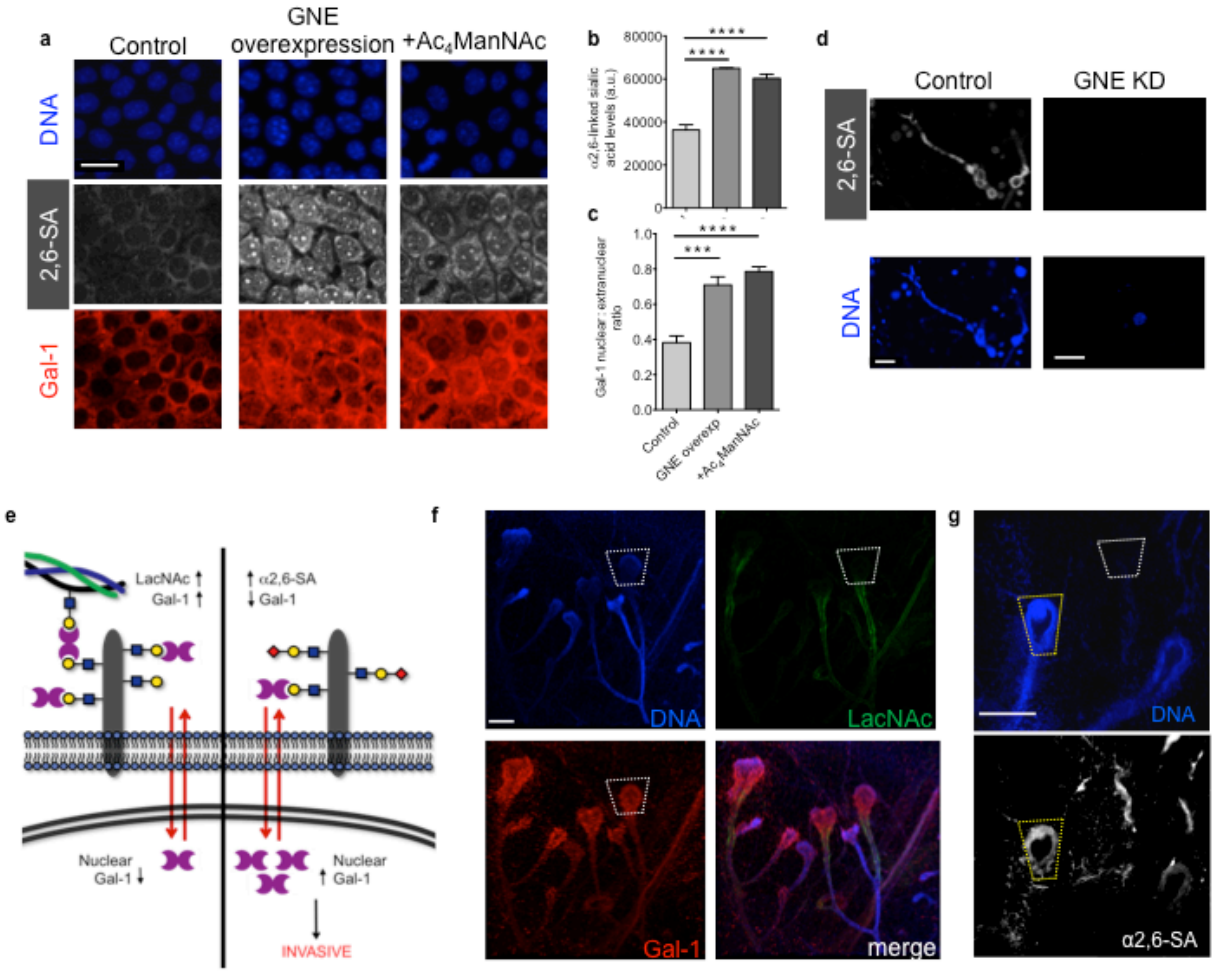


Figure 5.9. α2,6 SA regulates nuclear localization and mammary patterning by Gal-1. **a**, Fluorescence micrographs of EpH4 cells overexpressing GNE (middle panel), the rate limiting enzyme in sialic acid biosynthesis, or treated with the small molecule $Ac_4ManNAc$ (right panel), which is processed by the sialic acid salvage pathway. EpH4 cells were stained with *Sambucus nigra* lectin (SNA), which is specific for α2,6 linked SA , and with an α-Gal-1 antibody. Scale bar, 50 µm. **b**, Overexpression of GNE and $Ac_4ManNAc$ administration increases α2,6 linked SA levels in EpH4 cells. **c**, Quantification of Gal-1 nuclear:extranuclear ratio of EpH4 cells with varying levels of cell surface sialylation. **d**, Fluorescence micrographs of GNE KD EpH4 cells cultured in a 3D collagen I gel and stained with SNA and DAPI. Branching and invasion is absent in EpH4 cells lacking α2,6 SA residues. Scale bar, 50 µm. **e**, Model of extracellular glycan patterns regulating nuclear Gal-1 in mammary epithelial cells. In a LacNAc-rich environment containing intact ECM proteins, Gal-1 is mainly concentrated in the extracellular space (left). However, α2,6 sialylation of LacNAc structures causes Gal-1 to accumulate in the nucleus, resulting in an invasive phenotype (right). **f**, Immunofluorescence micrographs of murine mammary gland stained with *Erythrina Crystagalli* lectin (ECL) (top, right), which is specific for terminal LacNAc disaccharides, an α-Gal-1 antibody (bottom, left), and DAPI (top, left). LacNAc appears to line the ductal epithelia, whereas Gal-1 is heavily concentrated in the invasive end bud. Scale bar, 150 µm. **g**, Fluorescence micrographs of murine mammary gland stained with SNA (bottom) and DAPI (top) show high levels of α2,6 SA residues in the invasive

end bud of the mammary gland. Scale bar, 200 μm. For all bar graphs, error bars represent s.e.m. Statistical significance is given by ***$P < 0.001$; ****$P < 0.0001$.

Our data therefore point to a dynamic reciprocity between the glycan microenvironment and nuclear Gal-1 levels. In the case when the extracellular environment contains an excess of Gal-1 ligands, e.g. terminal LacNAc epitopes, the 'equilibrium' favors the accumulation of Gal-1 outside of the cell (Figure 5.9e). Here, the extracellular glycan environment acts as a molecular sink for Gal-1, trapping it. On the other hand, when Gal-1 is unable to bind extracellular glycan ligands, e.g. GAL-1(N46D) or due to presence of α2,6-SA residues, the 'equilibrium' shifts to a higher abundance of nuclear Gal-1, promoting epithelial invasiveness.

We sought to determine if this mechanism is relevant to, and operative during, *in vivo* mammary branching morphogenesis. Towards this end, 5 week-old fixed mammary gland whole mounts were stained with a FITC-*Erythrina Crystagalli* lectin (ECL), which recognizes unmodified terminal LacNAc residues,[23] and for Gal-1. Fluorescence microscopy displayed mutually exclusive signals for nuclear Gal-1 and extracellular LacNAc: nuclear Gal-1 was strongly present in the TEBs whereas LacNAc levels were highest in the ducts (Figure 5.9f). Sections of 5-week mammary gland ducts stained for LacNAc showed strong colocalization with the basement membrane and low levels of extracellular Gal-1 (Figure 5.10). We also confirmed our *in vivo* glycan findings with branching mammary epithelia clusters in culture (Figure 5.11): fixed branching mammary epithelia showed high levels of LacNAc, except at the invasive tips which showed little to no LacNAc. These tips, on the other hand, had high levels of nuclear Gal-1.

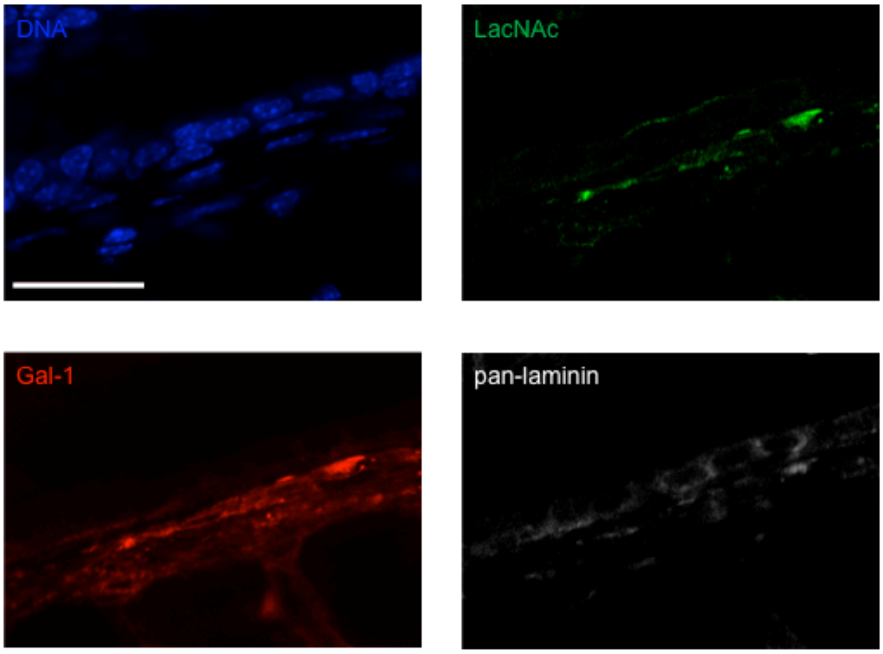


Figure 5.10. Gal-1 colocalizes with terminal LacNAc and laminin in mammary ductal epithelia. Immunofluorescence micrographs of a five-week murine mammary gland section showing

mammary ductal epithelia stained with DAPI (top, left), α-Gal-1 antibody (bottom, left), ECL (top, right), and α-pan-laminin antibody (bottom, right). Scale bar, 20 μm.

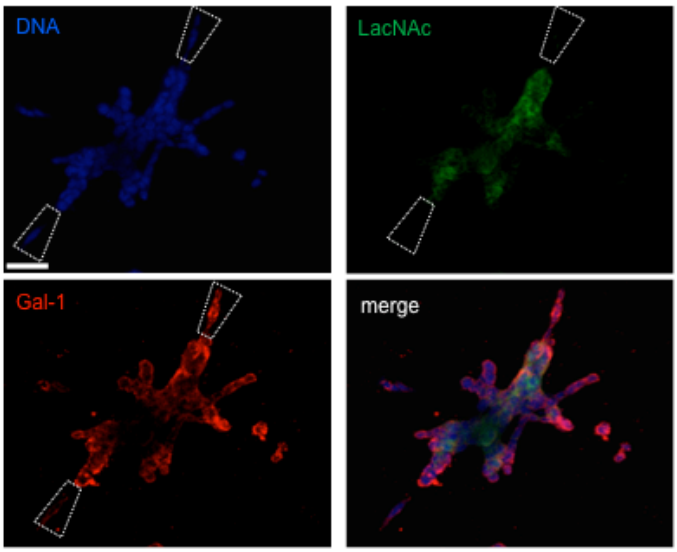


Figure 5.11. Localization of Gal-1 and terminal LacNAc in EpH4 cells cultured in 3D. Immunofluorescence micrographs of branching Eph4 cells cultured in 3D CL-I gels stained with DAPI (top, left), α-Gal-1 antibody (bottom, left) and ECA (top, right). Scale bar, 50 μm.

Consistent with our model, we also detected high levels of α2,6-SA in the terminal end buds *in vivo* and the branching mammary epithelia in 3D cultures (Figure 5.9g). This result is in agreement with data from a microarray set which showed higher gene expression for ST6Gal1, an enzyme that adds α2,6-SA and blocks Gal-1 binding, within end bud epithelia compared to ductal epithelia[24,25].

Conclusion

In summary, we have demonstrated that glycans of the mammary tissue microenvironment determine the invasive and branching behavior of mammary epithelia by regulating the nuclear localization of Gal-1. LacNAc, the cognate glycan ligand for Gal-1, acts as a sink to retain Gal-1 in the extracellular mileu. We found significant levels of unmodified LacNAc and low nuclear Gal-1 levels in the quiescent ductal epithelia of mammary glands. α2,6-siaylation increased nuclear Gal-1 levels by disrupting the binding of Gal-1 to glycoconjugates. We found significant levels of α2,6-SA and high nuclear Gal-1 levels in the proliferating epithelia at the invading edge of mammary end buds. The induction of invasion and branching by Gal-1 in the nucleus could be mediated by its positive regulation of Erk1/2, a mitogen activating protein kinase (MAPK), levels and activity[26] (Figure 5.12).

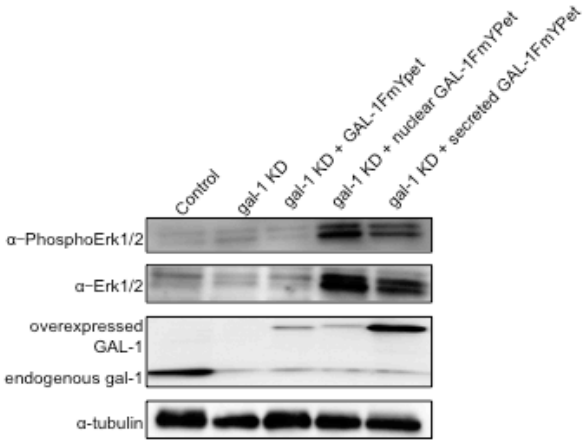


Figure 5.12. Nuclear Gal-1 enhances Erk1/2 activity. Immunoblot (left) showing levels of endogenous Gal-1, ectopically expressed GAL-1, Erk1/2 and phosphorylated Erk1/2 (from bottom to top) in control (scrambled shRNA) cells, Gal-1 KD cells, Gal-1 KD cells ectopically expressing GAL-1, Gal-1 KD cells ectopically expressing NLS-GAL-1, and GAL-1 KD cells ectopically expressing SEC-GAL-1 (from left to right). Tubulin levels indicate uniform loading.

In vitro studies have shown that nuclear Gal-1 is involved in pre-mRNA splicing[27]. We observed that Gal-1 localizes to the spliceosomes of mammary epithelia (Figure 5.13a), suggesting that it may be part of Gemin-4-containing spliceosomal complexes. A knockdown of Gemin-4 expression in mammary epithelia using shRNA (Figure 5.13b) resulted in an abrogation of branching: the cells remained alive and formed noninvasive spherical clusters, phenocopying Gal-1 KD cells (Figure 5.13c). This strongly suggests that the interaction of Gemin-4 and Gal-1 within the nucleus can modulate mammary epithelial invasiveness. This interaction could also be the mechanism by which nuclear Gal-1 regulates the levels of proteins such as MAPK and in turn promotes differential mammary end bud proliferation. Invasive breast cancer cells have high levels of α2,6-SA[3,28,29], which blocks Gal-1 binding, as well as an increase in Gal-1 protein levels[30,31]. Our results provide insight into this paradoxical finding. Translocation of Gal-1 to the nucleus of malignant breast cancer due to hypersialylation may provide a path to cancer cell invasion, which has implications for the future of Gal-1 specific inhibitors in breast cancer treatment.

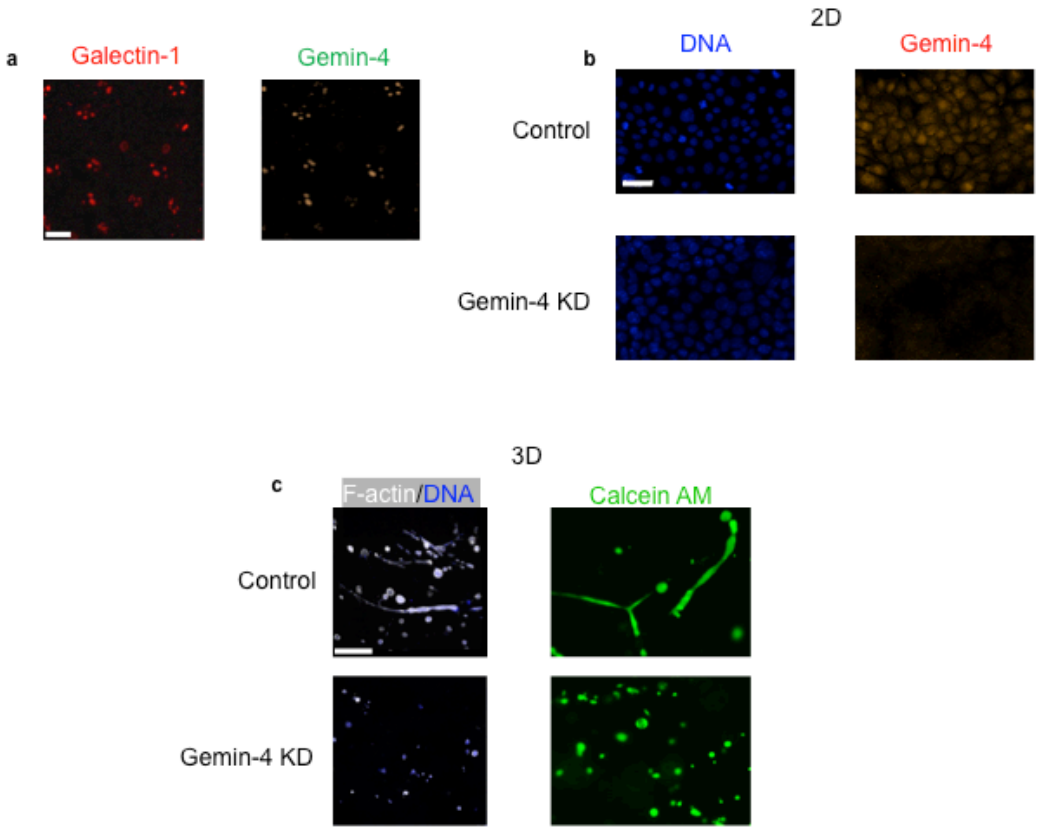


Figure 5.13. Association of Gal-1 and Gemin-4 in mammary epithelial nuclei. **a,** Immunofluorescence micrographs of EpH4 cells depleted of soluble nuclear chromatin fraction, cytoplasmic fraction and DNA, stained with an α-Gal-1 antibody and an α-Gemin-4 antibody. Scale bar, 200 µm. **b**, Immunofluoresence micrographs of EpH4 cells, control (top) and with Gemin-4 KD, cultured in 2D and stained with DAPI and α-Gemin-4 antibody. Scale bar, 50 µm. **c**, Immunofluorescence micrographs of control and Gemin-4 KD EpH4 cells cultured in 3D and stained with Phalloidin and DAPI (left) to assess branching morphology and Calcein AM vital dye (right) to assess viability. Scale bar, 200 µm.

Experimental Procedures

Immunofluorescence

Freshly dissected mammary glands were fixed with 4% paraformaldehyde for 15 minutes and then with Carnoy's fixative (75% ethanol and 25% glacial acetic acid) overnight. This was followed by serial dehydration and overnight treatment with xylene. Dehydrated tissues were processed for paraffin embedding and sectioning. H&E stained and unstained 5 mm tissue paraffin sections were generated by the UCSF Helen Diller Family Comprehensive Cancer Center Mouse Pathology Core. Deparaffinized sections were subjected to microwave-based antigen retrieval using citrate buffer (pH 6) for 10 min. The sections were then blocked with PBS containing 5% normal goat serum, 1% BSA and 0.05% Tween-20 for 1 hour at rt, followed by incubation with rabbit polyclonal α-Gal-1 antibody (Abcam ab25138), diluted 1:200, overnight at 4°C. All images were obtained using a laser scanning confocal microscope LSM710 (Zeiss).

Immunostaining of mammary whole-mounts

Mammary whole-mount of 5-week-old C57BL/6 mice was used. The mammary tissue was processed as previously described[11] with some modifications in order to perform immunostaining. The antigen retrieval was performed as described earlier. Then, the mammary whole-mount was incubated with Triton-X 100 (0.1%) for 25 min and blocked (5% normal goat serum + 1% BSA + 0.05% Tween-20) overnight at 4°C. This was followed by incubation with anti-heparanase antibody, diluted at 1:15, overnight at 4°C. The tissue was washed for 24 hours. The secondary antibody used was Alexa Fluor 633-conjugated goat anti-rabbit (1:100), incubated overnight. Finally, the tissue was washed for 24 hours, dehydrated and mounted in Permount (Fisher Scientific, PA). Images were obtained using a confocal microscope (LSM710, Zeiss) using 10X and 40X lens. The lambda scan mode in the confocal was utilized to determine the maximum emission of the autofluorescence and from the secondary antibody.

EpH4 cell culture and preparation of cell clusters

EpH4 cells were maintained in 1:1 DMEM/F12, supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA) and 50 μg/ml gentamycin (EpH4 media). The EpH4 cells were used between passage 10 and 20. EpH4 cells suspended in growth medium were plated in six-well polyHEMA-coated plates ($2x10^5$ cells per well) and incubated overnight at 37°C, yielding rounded clusters. Single cells were removed by differential centrifugation, and the final pellet was plated again in six-well plates overnight. Finally, cell clusters were centrifuged and resuspended in the desired amount of the medium described above.

Three dimensional (3D) collagen I (CL-1) gel

For 3D cell cultures, EpH4 cell clusters were embedded in CL-1 gels. Briefly, acid-soluble collagen (900 μl of a 5 mg/ml solution), Cellagen (AC-5, ICN, Koken, Tokyo, Japan) was gently mixed on ice with 112.5 μL 10X DMEM/F12, followed by 85 μL of 0.1 N NaOH and 375 μL of DMEM/F12. Around 250 clusters were plated in each well. Two layers of collagen were poured into each well: a basal layer consisting only of CL-1 and an upper layer containing

clusters of EpH4 cells. To allow solidification of the basal CL-1 layer, 75 μL of the CL-1 solution was poured into each well of a 48-well dish and incubated at 37 °C for 15 min. A second layer of 220 μL of CL-1 containing 250 EpH4 clusters was added to each well and placed immediately at 37°C. After gelation, 300 μL of chemically defined medium (DMEM/F-12 containing 1% insulin/transferrin/selenium, 1% penicilin/streptomycin) with 9 nM bFGF (Sigma) was added to each well and replaced every other day. Where indicated, Calcein AM dye (Life Technologies) was used to assess for cell viability. As exogenous treatments, we used 25 μM Ac_4ManNAc (controls were treated with an appropriate amount of vehicle DMSO). In experiments with exogenous Gal-1, human recombinant Gal-1, either unconjugated, or conjugated with Alexa Fluor 633 using a Protein labeling kit (Life Technologies), was used.

Analysis of branching morphogenesis

The branching phenotype of EpH4 clusters embedded in collagen I gel was determined after a 5-day culture period. The branching phenotype was defined as a cell cluster having at least two processes extending from its central body. Quantification of EpH4 cell branching was carried out by counting the number of branching clusters in each well. In addition, we analyzed the number of branches in each cluster. All experiments were repeated at least three times.

shRNA and constructs of expression

To transduce FLAG-tagged human GAL-1F, NLS-GAL-1F (with nuclear localization signal sequence, PHPPKRLRSDPDAC, from Gemin-4 as predicted by cNLSMapper[32]), NES-GAL-1F (with nuclear exclusion signal sequence, INQMFSVQLSL, from Staufen-2[33]), SEC-GAL-1 (with signal peptide sequence, MYSMQLASCVTLTLVLLVNS, from Interleukin-2), N46D GAL-1 and myc-tagged GNE, each cDNA was made by PCR, confirmed by sequencing and ligated into pLenti-EF1α-puro, generated in our laboratory. Lentivirus plasmids containing shRNA (Mission shRNA; Sigma, St Louis, USA) against mouse Gal-1, Gemin-4 and GNE or lentivirus plasmids containing GAL-1 or variant constructs were transfected into 293FT cells using FuGene6 (Roche, Basel, Switzerland). Transfected cells were cultured in DMEM containing 5% FBS, 100 U/ml penicillin and 100 μg/ml streptomycin. Culture media was replaced after 24 hours with fresh media. Forty-eight hours later, recombinant lentivirus was concentrated from filtered culture media (0.45 μm filters) by ultracentrifugation at 100,000 g for 90 minutes (SW41Ti rotor; Beckman Coulter, Brea, USA). To transduce EpH4 cells, 1.0×10^5 cells were plated in each well of a six-well plate, infected with the lentivirus, treated with polybrene for 30 minutes and selected by adding 5 μg/ml puromycin to growth medium for 4 days. Lentivirus with scrambled sequence was used as an shRNA control. Template plasmids for Ypet and mCherry were purchased from Addgene (http://www.addgene.org) and alanine 206 mutated to lysine (monomeric). Monomeric Ypet or mCherry were fused, respectively, with the C terminus of GAL-1F or its variants by PCR. All the sequences were confirmed by sequencing.

Western blotting

To check knockdown and overexpression, EpH4 cells were cultured in EpH4 media or in CL-1 gel (3 mg/ml) for 48 hours in serum-free media (DMEM/F-12 containing 1% insulin/transferrin/selenium, 1% penicilin/streptomycin) with 9 nM bFGF (Sigma). For protein

isolation, media was removed, and the cells were washed with PBS and then lysed with a buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl, 0.5% Tween with protease and phosphatase inhibitor cocktails (EMD Millipore, Billerica, USA). Protein concentration was determined using the BCA Protein Assay kit (Thermo Scientific, Waltham, USA). Protein samples were mixed with Laemmli sample buffer and heated at 95°C for 5 minutes. Samples were loaded into a pre-cast 4-20% tris-glycine polyacrylamide gel (Invitrogen) using the NOVEX system (Invitrogen). Resolved proteins were transferred to nitrocellulose membrane (Whatman, Maidstone, UK) followed by blocking in PBS, 0.05% Tween-20 with 3% BSA for 1 hour at room temperature. Membranes were incubated overnight at 4°C in 5% BSA, 0.1% Tween-20 in PBS containing α-Gal-1 antibody (Abcam), anti-α-tubulin antibody (Sigma), anti-Erk1/2 antibody and anti-PhosphoErk1/2 antibody (Cell signaling). Primary antibodies were detected with the Pierce SuperSignal detection kit (Rockford, IL) and signal was captured with the FluorChem 8900 analysis system (Alpha Innotech, San Leandro, CA).

Glycopolymer treatment

Soluble lactosyl and cellobiosyl GPs were synthesized as previously described.[19] Briefly, polymerizations were carried out using standard Schlenk techniques. CTA **2.21** (4.09 mg, 4.55 μmol), lactosyl ethyl acrylamide **2.17** or cellobiosyl ethyl acrylamide **2.20** (0.100 g, 228 μmol), 4,4′-azobis(4-cyanovaleric acid) (319 μg, 1.14 μmol), DMF (200 μl) and H_2O (800 μl) were added to the Schlenk tube. The solution was degassed by sparging with N_2 for 30 min. The Schlenk flask was placed in a 70 ^oC bath to start the polymerization. After 16 h, the reactions were stopped by exposure to oxygen. The conversion for the glycosyl monomer was calculated by comparing the integrals of the alkene proton peaks of the monomer (5.7 ppm, 1H) and those of the backbone methylene protons on the glycopolymer (1.87- 1.00 ppm, 2H). Glycopolymers were purified by dialysis against H_2O for 72 h and lyophilized.

Lactosyl GP. According to the above procedures, **lactosyl GP** was obtained as a light yellow solid (90.2 mg, 89%). Conversion (NMR) = quant. PDI (SEC) = 1.1. ^1H NMR (500 MHz, D_2O): δ 4.61 – 4.33 (m, 100H), 4.12 – 3.20 (m, 800H), 2.33 – 1.87 (m, 47H), 1.87 – 1.38 (m, 77H), 1.10 (s, 12H).

Cellobiosyl GP. Cellobiosyl GP was prepared using the above procedure to yield a light yellow solid (87.9 mg, 87%). Conversion (NMR) = quant. PDI (SEC) = 1.2. ^1H NMR (500 MHz, D_2O): δ 4.65 – 4.36 (m, 100H), 4.15 – 3.12 (m, 800H), 2.35 – 1.87 (m, 41H), 1.87 – 1.38 (m, 73H), 1.38 – 1.09 (m, 26H).

EpH4 cells cultured on top or embedded within CL-1 gels were incubated with lactosyl GP, cellobiosyl GP, or no GP for 1- 3 days. After 1 day, GAL-1 expressing EpH4 cells on top of CL-1 gels were imaged using confocal fluorescence microscopy. After 3 days, EpH4 cells embedded within a CL-1 were washed extensively and then treated with α-Gal-1 antibody (Abcam), phalloidin, and DAPI. The cells were then imaged using confocal fluorescence microscopy.

Visualization of intranuclear bodies

For visualization of nuclear matrix proteins, extraction of soluble nuclear and cytoplasmic fractions was performed as previously described[34]. After washing with PBS, cells were extracted for 3 min with ice-cold cytoskeletal (CSK) buffer (10 mM PIPES, 300 mM sucrose, 100 mM NaCl, 3 mM MgCl2, 0.5% Triton X-100, pH 6.8) with protease inhibitors (Roche). Next, DNA was digested by DNAse I (Boehringer, Indianapolis, IN) in CSK buffer, 50 mM NaCl, for 1 h at 32 °C. Chromatin was then washed out with CSK buffer. Chromatin removal was confirmed with DAPI staining. Cells were then washed with 0.25 M ammonium sulfate in CSK followed by a 2M NaCl wash. Finally, cell were washed three times with PBS, fixed in 4% ultrapure paraformaldehyde as described above, immunostained (for endogenous Gal-1 and Gemin-4), and imaged with a laser scanning confocal microscope (LSM710, Zeiss) using 40X water immersion lens.

Statistics

Data were analyzed using the unpaired student's *t* test and error bars indicate the standard error of the mean.

References

(1) Feta, J. E.; Werb, Z.; Bissell, M. J. *Breast Cancer Res.* **2004**, *6*, 1-11.

(2) Lu, P.; Werb, Z. *Science* **2008**, *332*, 1506-1509.

(3) Julien, S.; Adriaenssens, E.; Ottenberg, K.; Furlan, A.; Courtand, G.; Vercoutter-Edouart, A. S.; Hanisch, F. G.; Delannoy, P.; Le Bourhis, X. *Glycobiology* **2006**, *16*, 54-64.

(4) Sahab, Z. J.; Man, Y. G.; Semaan, S. M.; Newcomer, R. G.; Byers, S. W.; Sang, Q. X. *Clin. Exp. Metastasis* **2010**, *27*, 493-503.

(5) Radisky, D.C.; Stallings-Mann, M.; Hirai, Y.; Bissell, M. J. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 228-234.

(6) Barondes, S. H.; Cooper, D. N.; Gitt, M. A.; Leffler, H. *J. Biol. Chem.* **1994**, *269*, 20807-20810.

(7) Camby, I.; Le Mercier, M.; Lefranc, F.; Kiss, R. *Glycobiology* **2006**, *16*, 137R-157-R.

(8) Beliveau, A.; Mott, J. D.; Lo, A.; Chen, E. I.; Koller, A. A.; Yaswen, P.; Muschler, J.; Bissell, M. J. *Genes Dev.* **2010**, *24*, 2800-2811.

(9) Lu, P.; Takai, K.; Weaver, V. M.; Werb, Z. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a005058.

(10) Bissell, M. J.; Hines, W. C. *Nat. Med.* **2011**, *17*, 320-329.

(11) Mori, H.; Borowsky, A. D.; Bhat, R.; Ghajar, C. M.; Seiki, M.; Bissell, M. J. *Am. J. Pathol.* **2012**, *180*, 2249-2256.

(12) Reichmann, E.; Ball, R.; Groner, B.; Friis, R. R. *J. Cell Biol.* **1989**, *108*, 1127-1138.

(13) Montesano, R.; Schaller, G.; Orci, L. *Cell* **1991**, *66*, 697-711.

(14) Hirai, Y.; Lochter, A.; Galosy, S.; Koshida, S.; Niwa, S.; Bissell, M. J. *J. Cell Biol.* **1998**, *140*, 159-169.

(15) Mori, H.; Lo, A. T.; Inman, J. L.; Alcaraz, J.; Ghajar, C. M.; Mott, J. D.; Nelson, C. M.; Chen, C. S.; Zhang, H.; Bascom, J. L.; Seiki. M.; Bissell, M. J. *Development* **2013**, *140*, 343-352.

(16) Cooper, D. N.; Barondes, S. H. *J. Cell Biol.* **1990**, *110*, 1681-1691.

(17) Rizqiawan, A.; Tobiume, K.; Okui, G.; Yamamoto, K.; Shigeishi, H.; Ono, S.; Shimasue, H.; Takechi, M.; Higashikawa, K.; Kamata, N. *Biochem. Biophys. Res. Commun.* **2013**, *441*, 904-910.

(18) Voss, P. G.; Gray, R. M.; Dickey, S. W.; Wang, W.; Park, J. W.; Kasai, K.; Hirabayashi, J.; Patterson, R. J.; Wang, J. L. *Arch. Biochem. Biophys.* **2008**, *478*, 18-25.

(19) Belardi, B.; O'Donoghue, G. P.; Smith, A. W.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9549-9552.

(20) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E.; Yagi, F.; Kasai, K.-ichi *Biochim. Biophys. Acta, Gen. Subj.* **2002**, *1572*, 232-254.

(21) Stowell, S. R.; Arthur, C. M.; Mehta, P.; Slanina, K. A.; Blixt, O.; Leffler, H.; Smith, D. F.; Cummings, R. D. *J. Biol. Chem.* **2008**, *283*, 10109-10123.

(22) Shibuya, N.; Goldstein, I. J.; Broekaert, W. F.; Nsimba-Lubaki, M.; Peeters, B.; Peumans, W. J. *J. Biol. Chem.* **1987**, *262*, 1596-1601.

(23) Wu, A. M.; Wu, J. H.; Tsai, M. S.; Yang, Z.; Sharon, N.; Herp, A. *Glycoconj. J.* **2007**, *24*, 591-604.

(24) Amano, M.; Galvan, M.; He, J.; Baum, L. G. *J. Biol. Chem.* **2003**, *278*, 7469-7475.

(25) Kouros-Mehr, H.; Werb, Z. *Dev. Dyn.* **2006**, *235*, 3404-3412.

(26) Fata, J. E.; Mori, H.; Ewald, A. J.; Zhang, H.; Yao, E.; Werb, Z.; Bissell, M. J. *Dev. Biol.* **2007**, *306*, 193-207.

(27) Vyakarnam, A.; Dagher, S. F.; Wang, J. L.; Patterson, R. J. *Mol. Cell. Biol.* **1997**, *17*, 4730-4737.

(28) Soares, R., Marinho, A. & Schmitt, F. *Pathol. Res. Pract.* **1996**, *192*, 1181-1186.

(29) Recchi, M. A.; Harduin-Lepers, A.; Boilly-Marer, Y.; Verbert, A.; Delannoy, P. *Glycoconj. J.* **1998**, *15*, 19-27.

(30) Dalotto-Moreno, T.; Croci, D. O.; Cerliani, J. P.; Martinez-Allo, V. C.; Dergan-Dylon, S.; Méndez-Huergo, S. P.; Stupirski, J. C.; Mazal, D.; Osinaga, E.; Toscano, M. A.; Sundblad, V.; Rabinovich, G. A.; Salatino, M. *Cancer Res.* **2013**, *73*, 1107-1117.

(31) Kreunin, P.; Yoo, C.; Urquidi, V.; Lubman, D. M.; Goodison, S. *Proteomics* **2007**, *7*, 299-312.

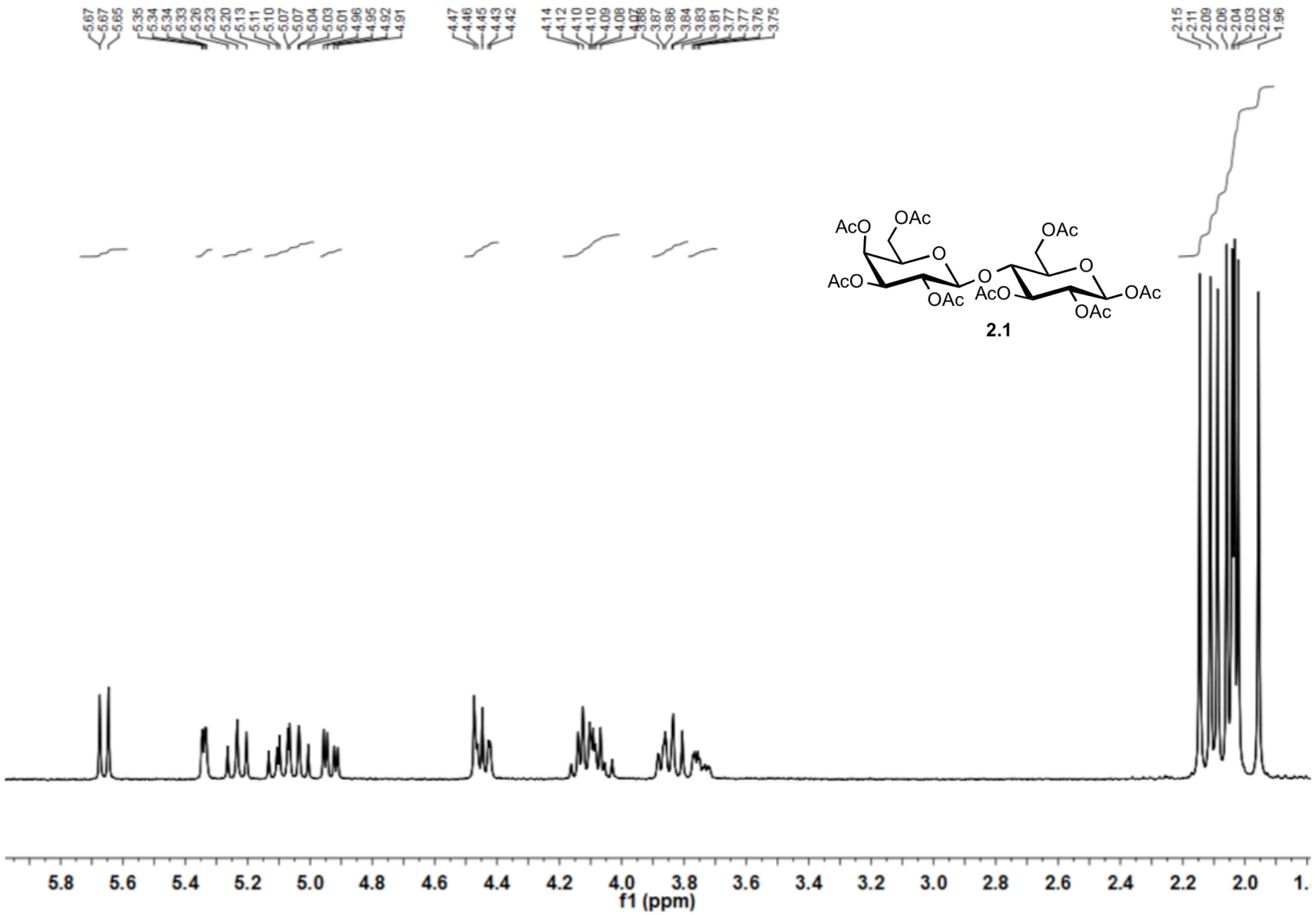
(32) Kosugi, S.; Hasebe, M.; Matsumura, N.; Takashima, H.; Miyamoto-Sato, E.; Tomita, M.; Yanagawa, H. *J. Biol. Chem.* **2009**, *284*, 478-485.

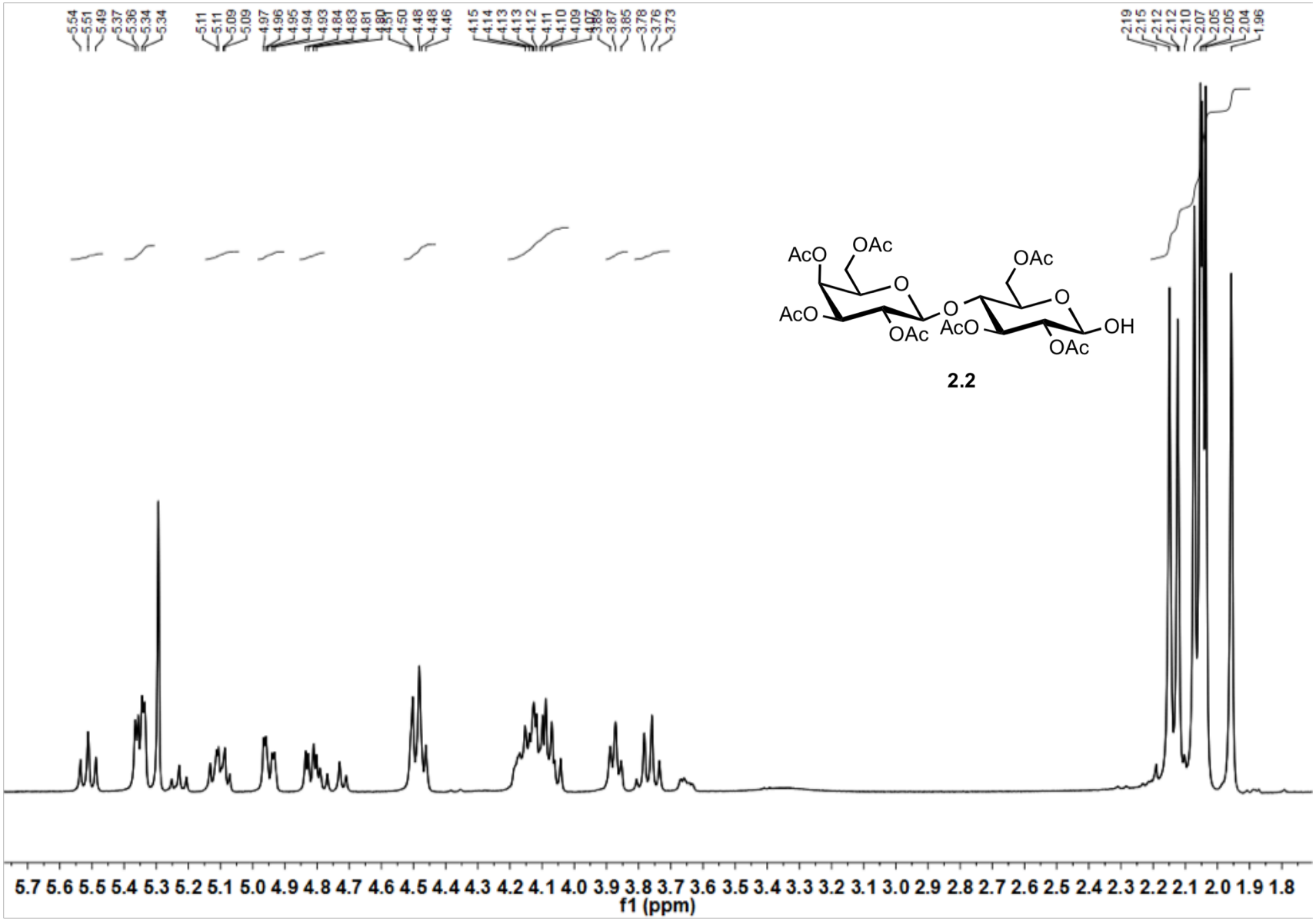
(33) Miki, T.; Yoneda, Y. *J. Biol. Chem.* **2004**, *279*, 47473-47479.

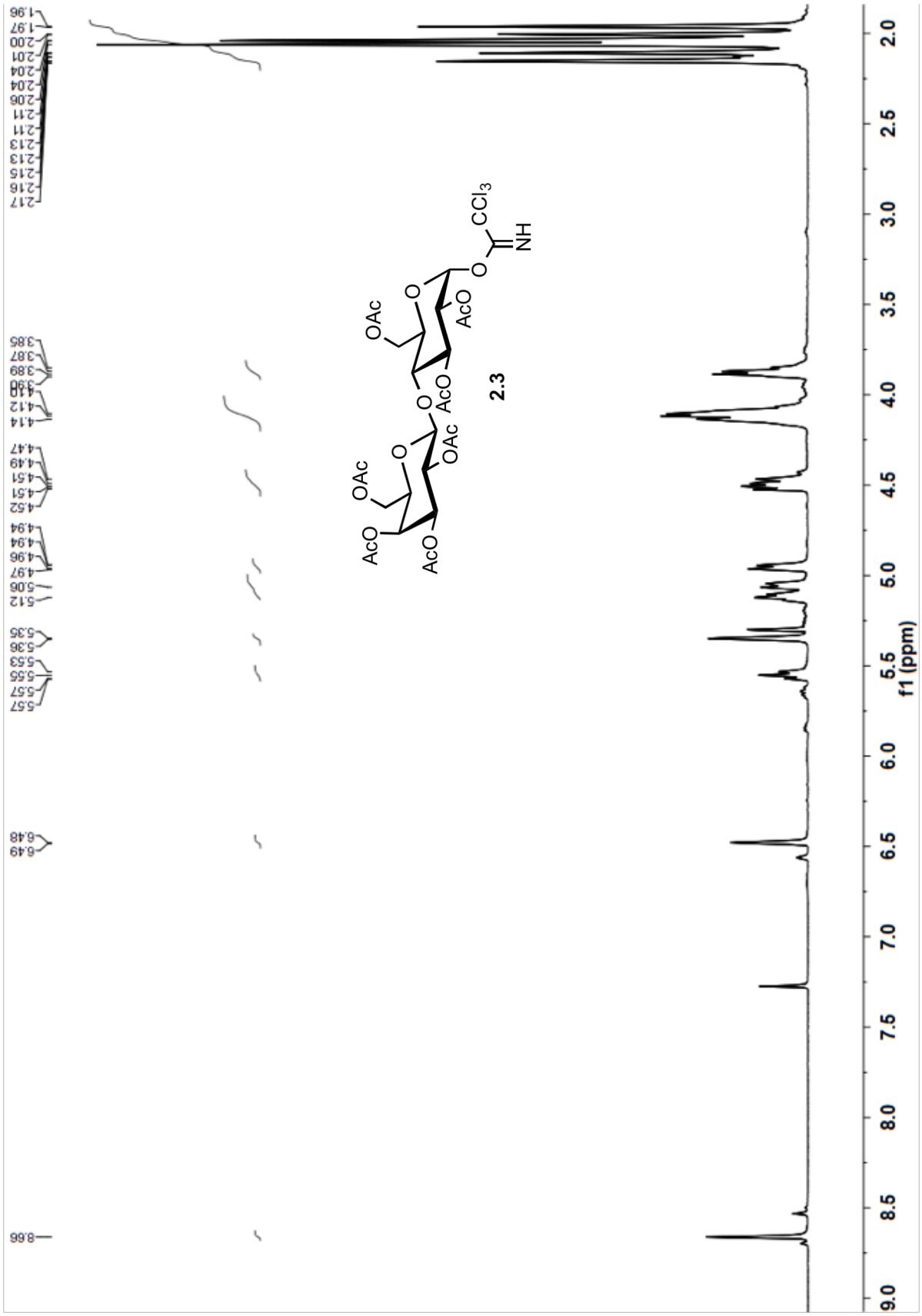
(34) Nalepa, G.; Harper, J. W. *Cell Motil. Cytoskeleton* **2004**, *59*, 94-108.

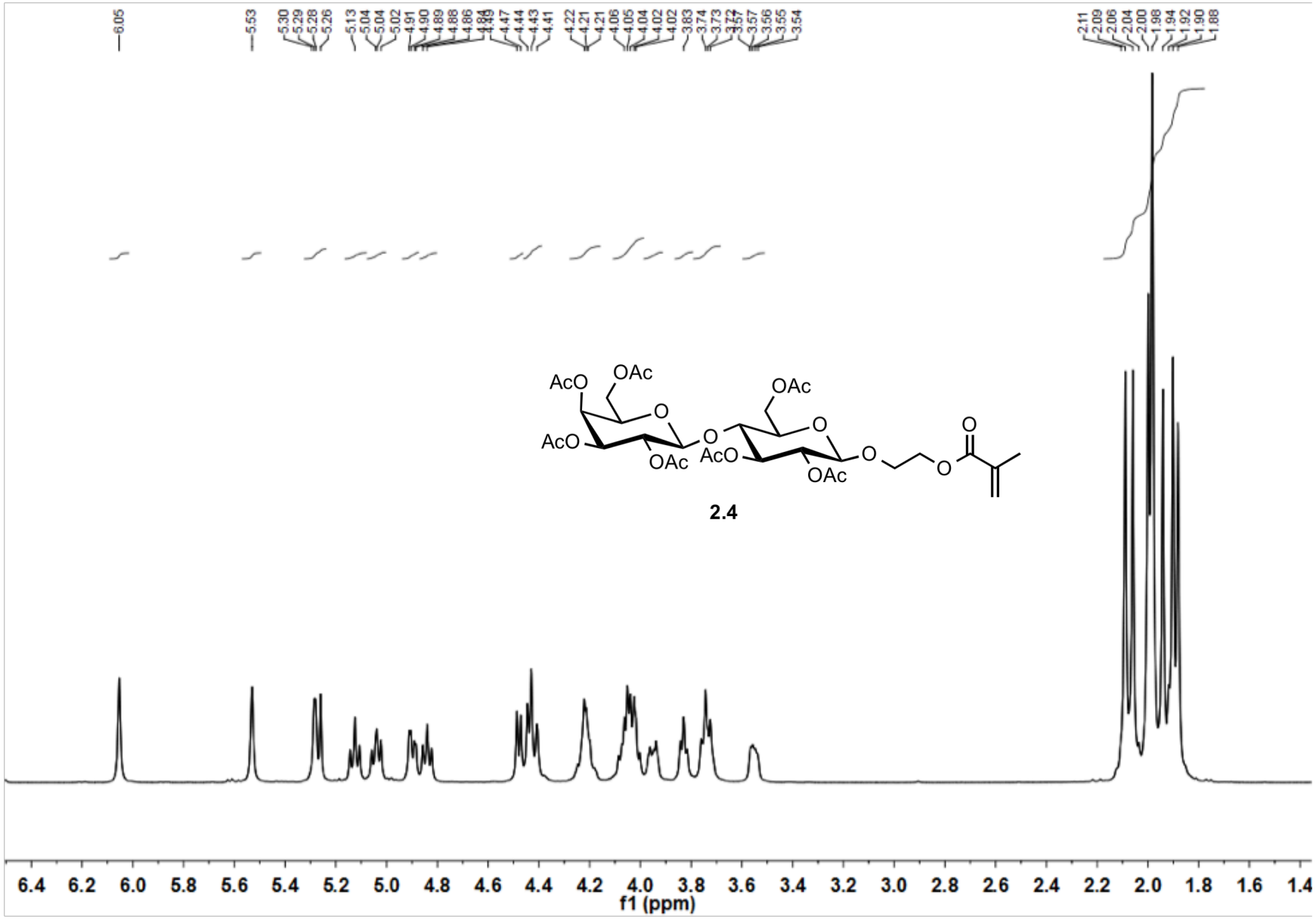
Appendix

^{1}H, ^{13}C, ^{31}P, and ^{19}F NMR and UV-Vis Spectra, GPC Traces

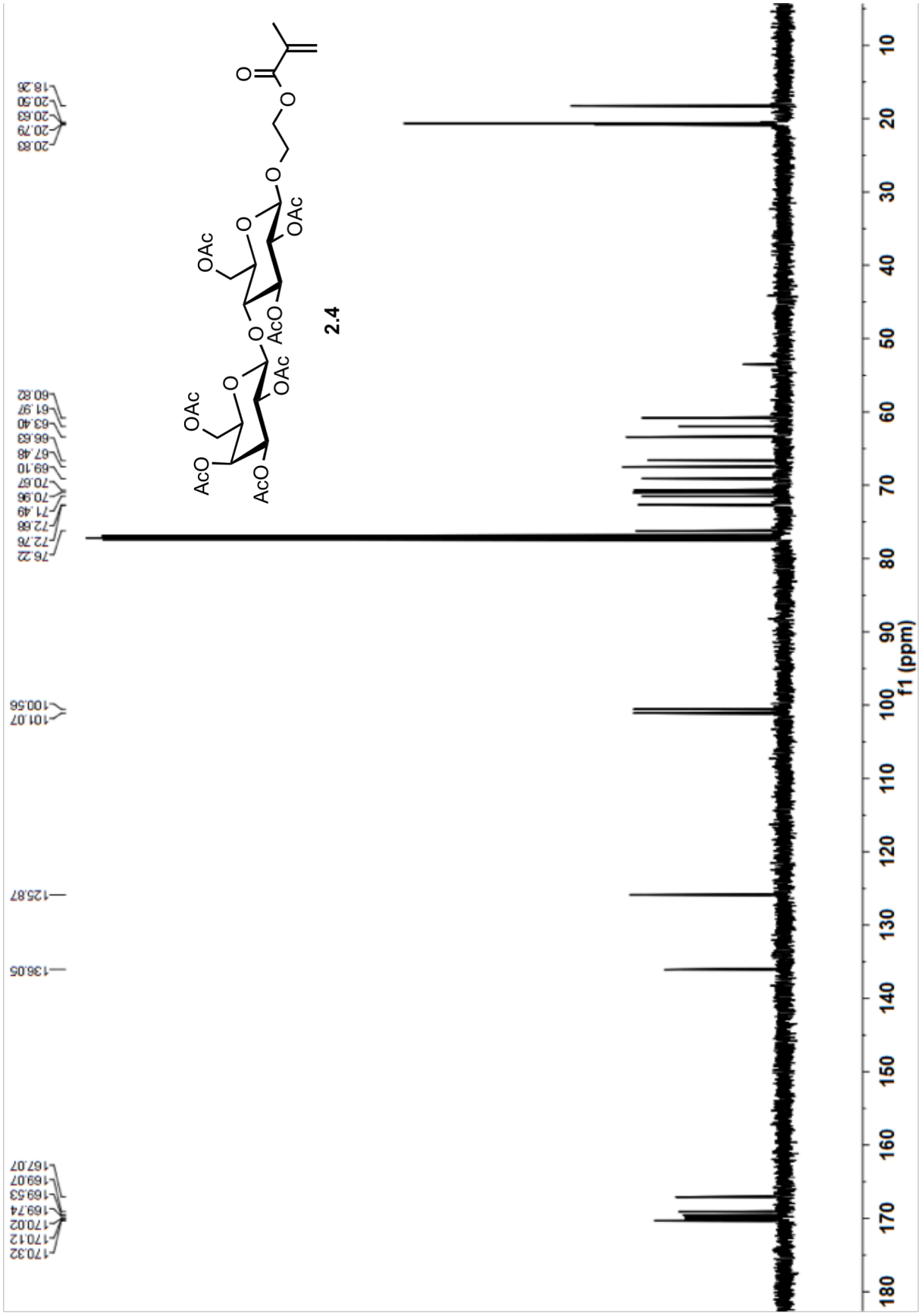

5.67
5.67
5.65
5.35
5.34
5.34
5.33
5.26
5.23
5.20
5.13
5.11
5.10
5.07
5.07
5.04
5.03
5.01
4.96
4.95
4.92
4.91
4.47
4.46
4.45
4.43
4.42
4.14
4.12
4.10
4.10
4.09
4.08
3.87
3.86
3.84
3.83
3.81
3.77
3.77
3.76
3.75
2.15
2.11
2.09
2.06
2.04
2.03
2.02
1.96
AcO
OAc
AcO
OAc
O
O
AcO
OAc
OAc
O
OAc
OAc
2.1
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.
f1 (ppm)

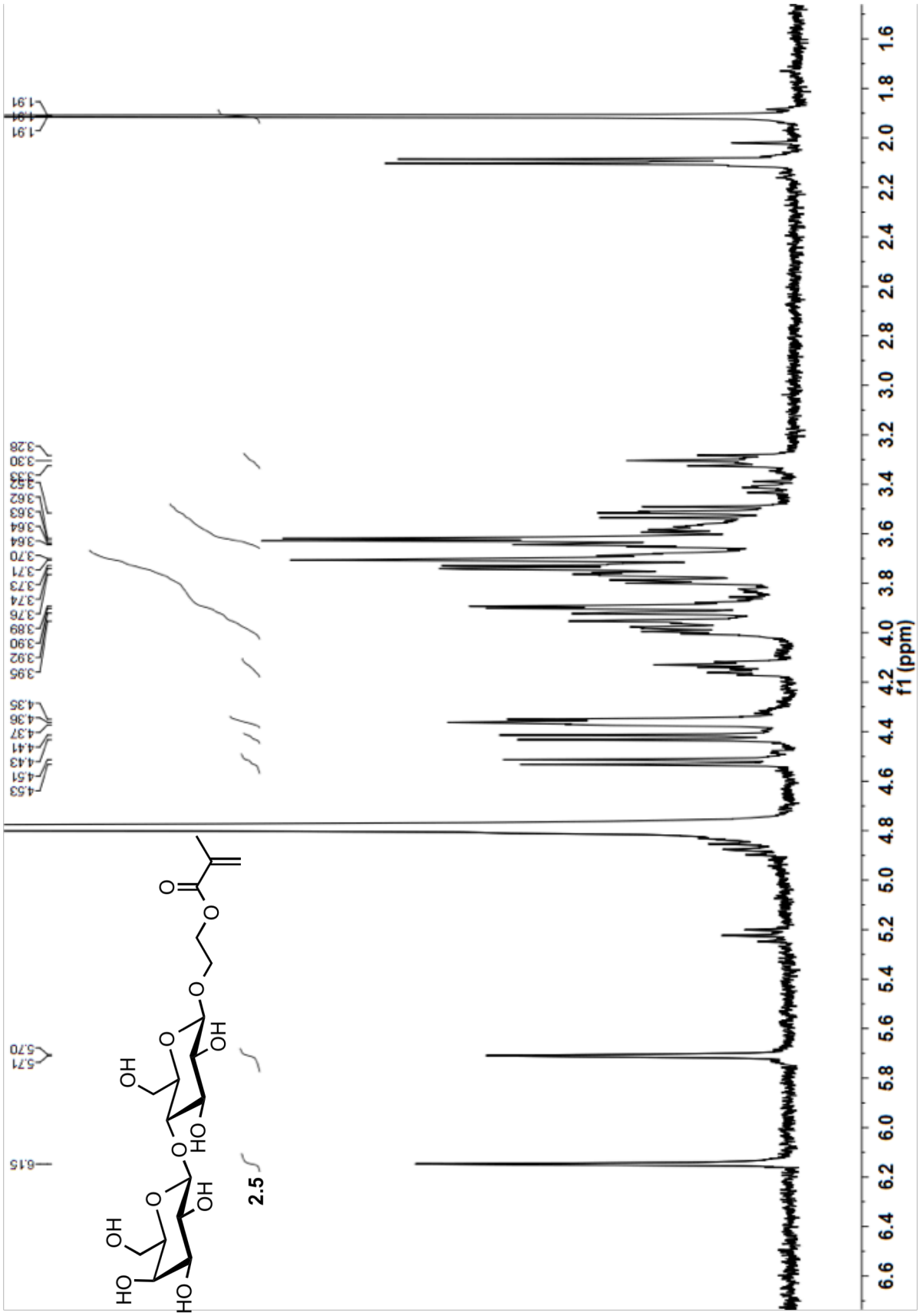

5.54
5.51
5.49
5.37
5.36
5.34
5.34
5.11
5.11
5.09
5.09
4.97
4.96
4.95
4.94
4.93
4.84
4.83
4.81
4.80
4.51
4.50
4.48
4.48
4.46
4.15
4.14
4.13
4.13
4.12
4.11
4.10
4.09
4.07
3.89
3.87
3.85
3.78
3.76
3.73
2.19
2.15
2.12
2.12
2.10
2.07
2.05
2.05
2.04
1.96
AcO
OAc
O
AcO
OAc
O
OAc
O
AcO
OAc
OH
2.2
5.7 5.6 5.5 5.4 5.3 5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1 3.0 2.9 2.8 2.7 2.6 2.5 2.4 2.3 2.2 2.1 2.0 1.9 1.8
f1 (ppm)

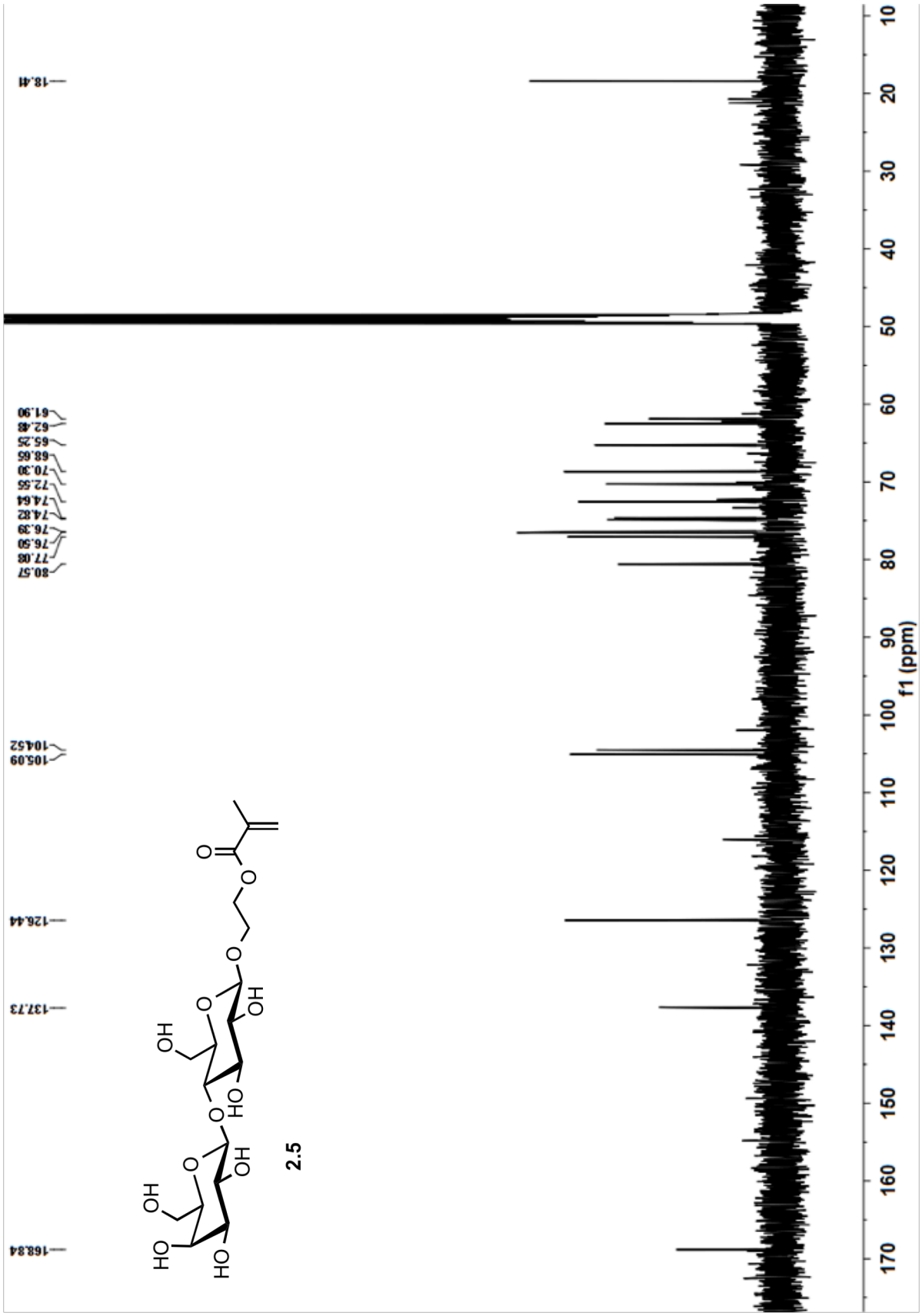

8.66
6.49
6.48
5.57
5.55
5.53
5.36
5.35
5.12
5.06
4.97
4.96
4.94
4.94
4.52
4.51
4.51
4.49
4.47
4.14
4.12
4.10
3.90
3.89
3.87
3.85
2.17
2.16
2.15
2.13
2.13
2.11
2.11
2.06
2.04
2.04
2.01
2.00
1.97
1.96
AcO
OAc
O
AcO
OAc
O
AcO
OAc
O
O
CCl3
NH
2.3
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
f1 (ppm)

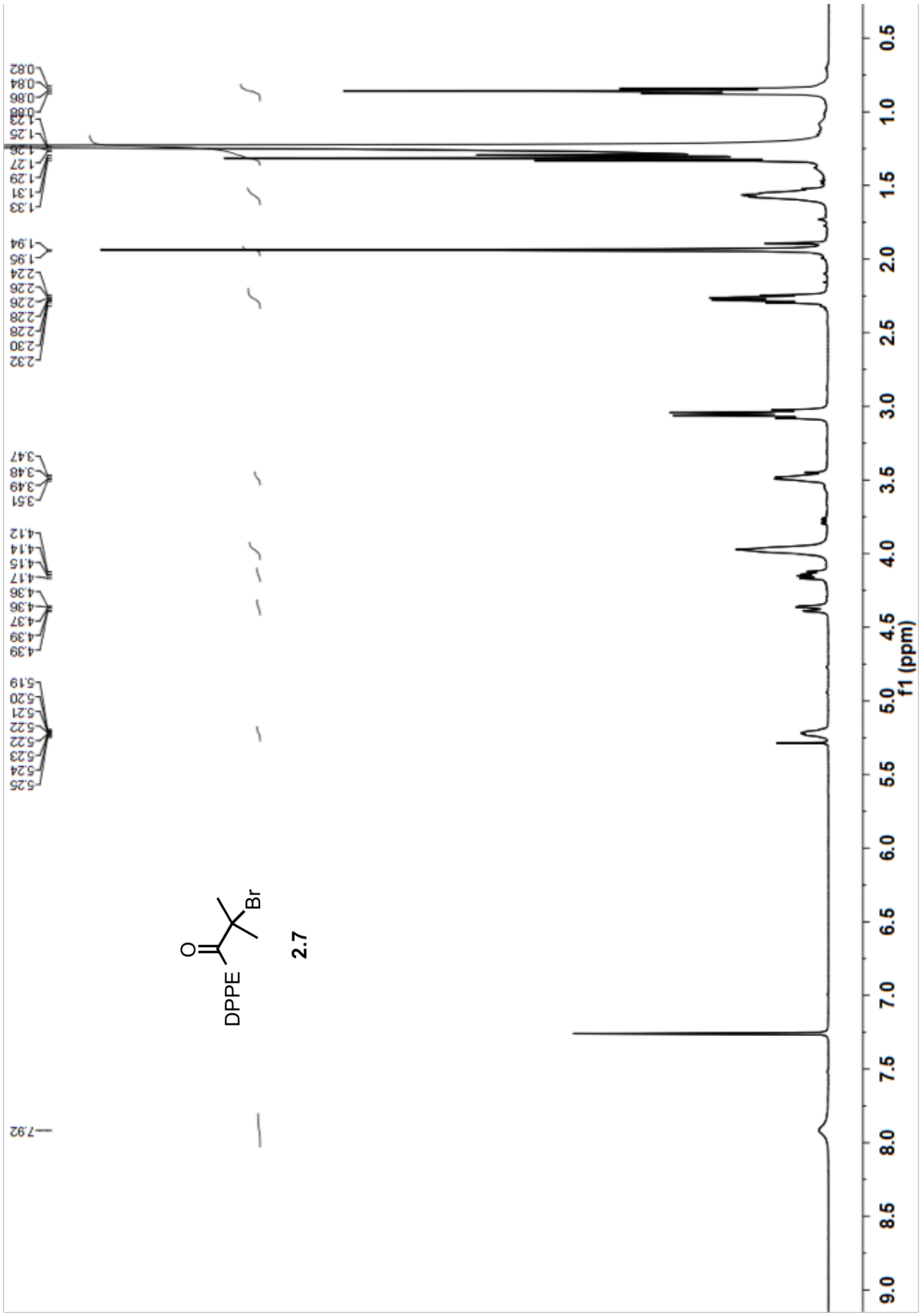

6.05
5.53
5.30
5.29
5.28
5.26
5.13
5.04
5.04
5.02
4.91
4.90
4.89
4.88
4.86
4.84
4.47
4.44
4.43
4.41
4.22
4.21
4.21
4.06
4.05
4.04
4.02
4.02
3.83
3.74
3.73
3.57
3.57
3.56
3.55
3.54
2.11
2.09
2.06
2.04
2.00
1.98
1.94
1.92
1.90
1.88
AcO
OAc
AcO
OAc
O
O
OAc
AcO
OAc
O
O
O
O
6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 3.2 3.0 2.8 2.6 2.4 2.2 2.0 1.8 1.6 1.4
f1 (ppm)

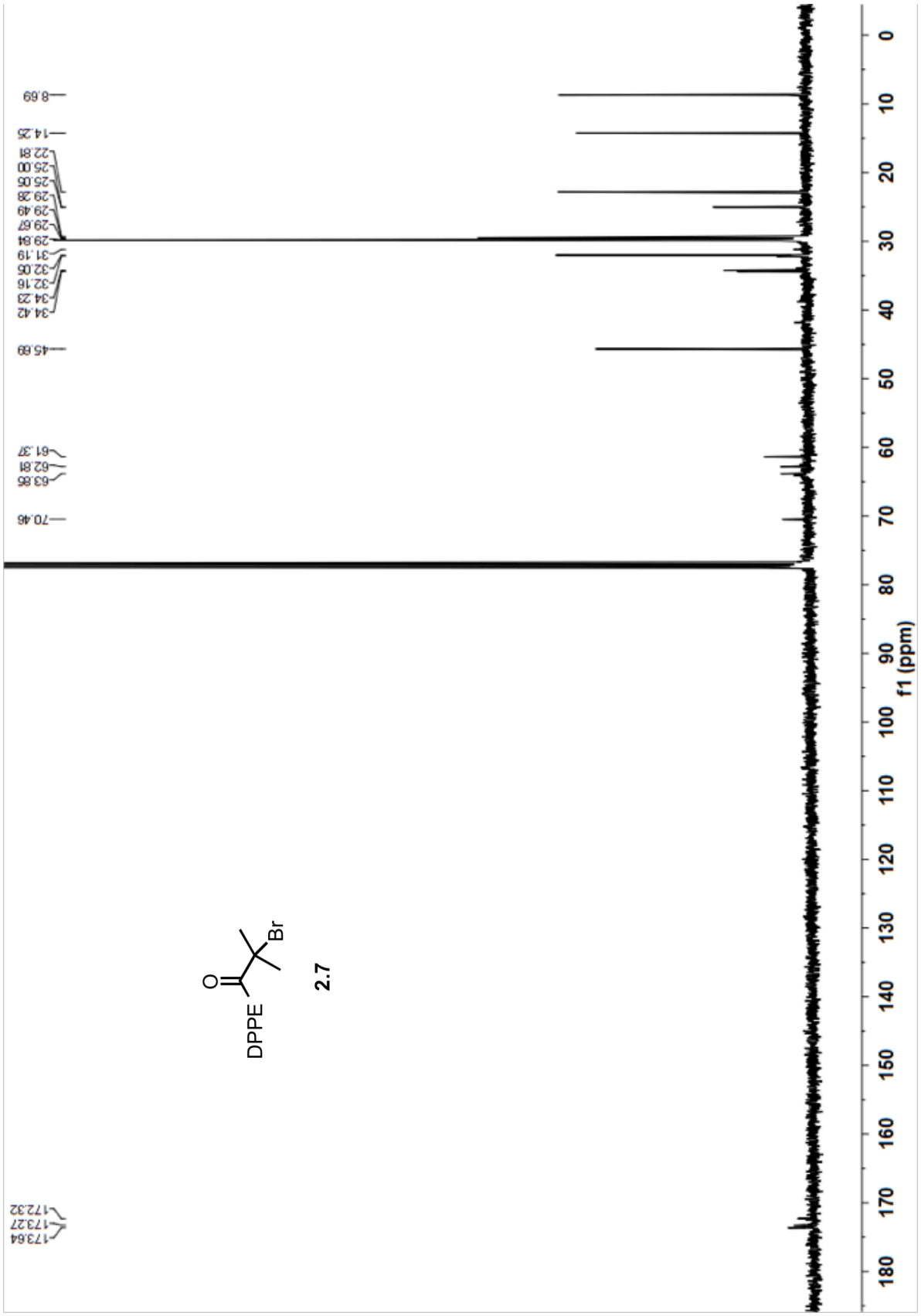
2.4

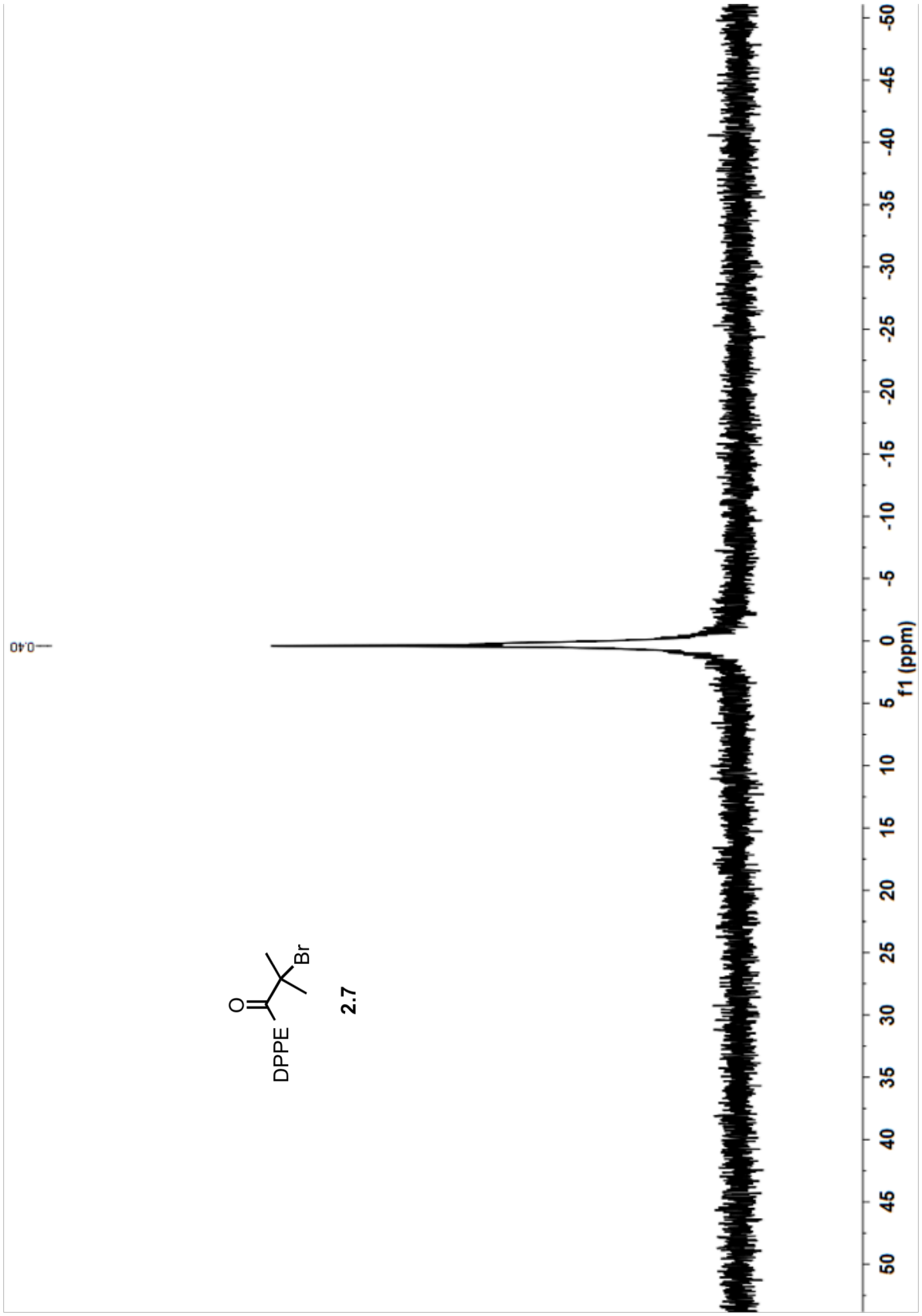

2.4
f1 (ppm)

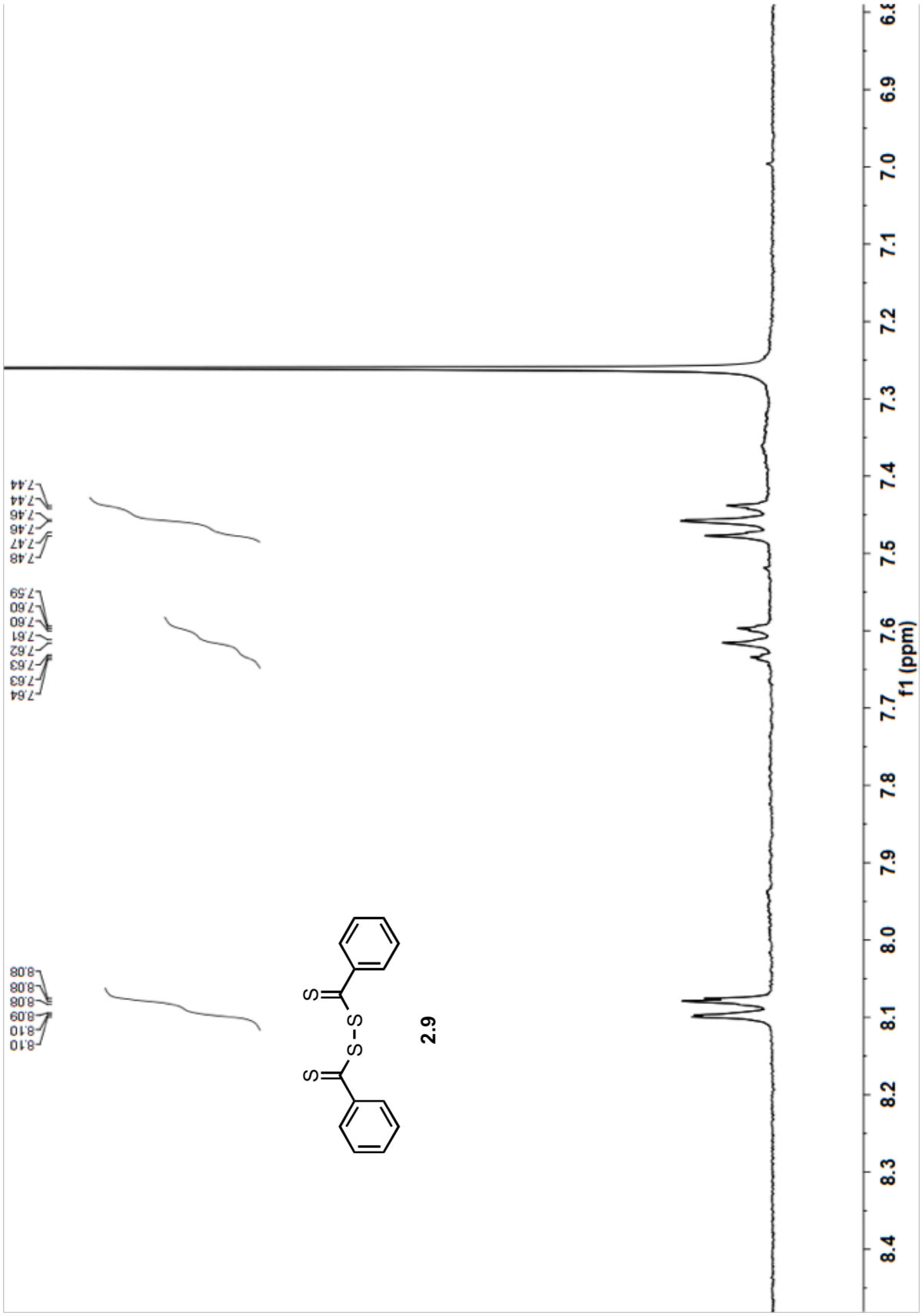

1.91
1.94
1.91
3.28
3.30
3.33
3.52
3.62
3.63
3.64
3.64
3.70
3.71
3.73
3.74
3.76
3.89
3.90
3.92
3.95
4.35
4.36
4.37
4.41
4.43
4.51
4.53
5.71
5.70
6.15
2.5
6.6
6.4
6.2
6.0
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
f1 (ppm)
HO
OH
O

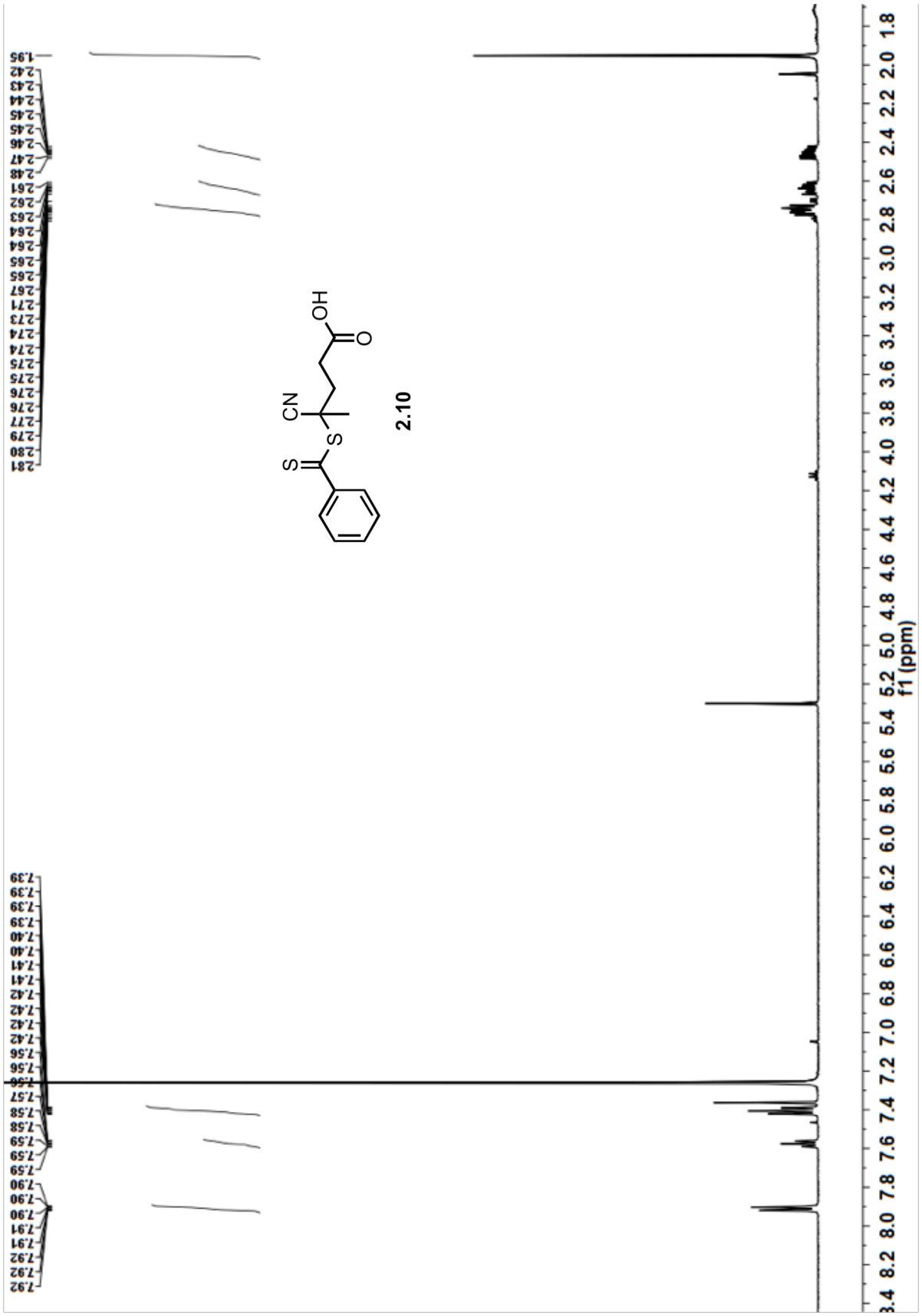

18.41
61.90
62.48
65.25
68.65
70.30
72.55
74.64
74.82
76.39
76.50
77.08
80.57
105.09
104.52
126.44
137.73
168.84
2.5
f1 (ppm)

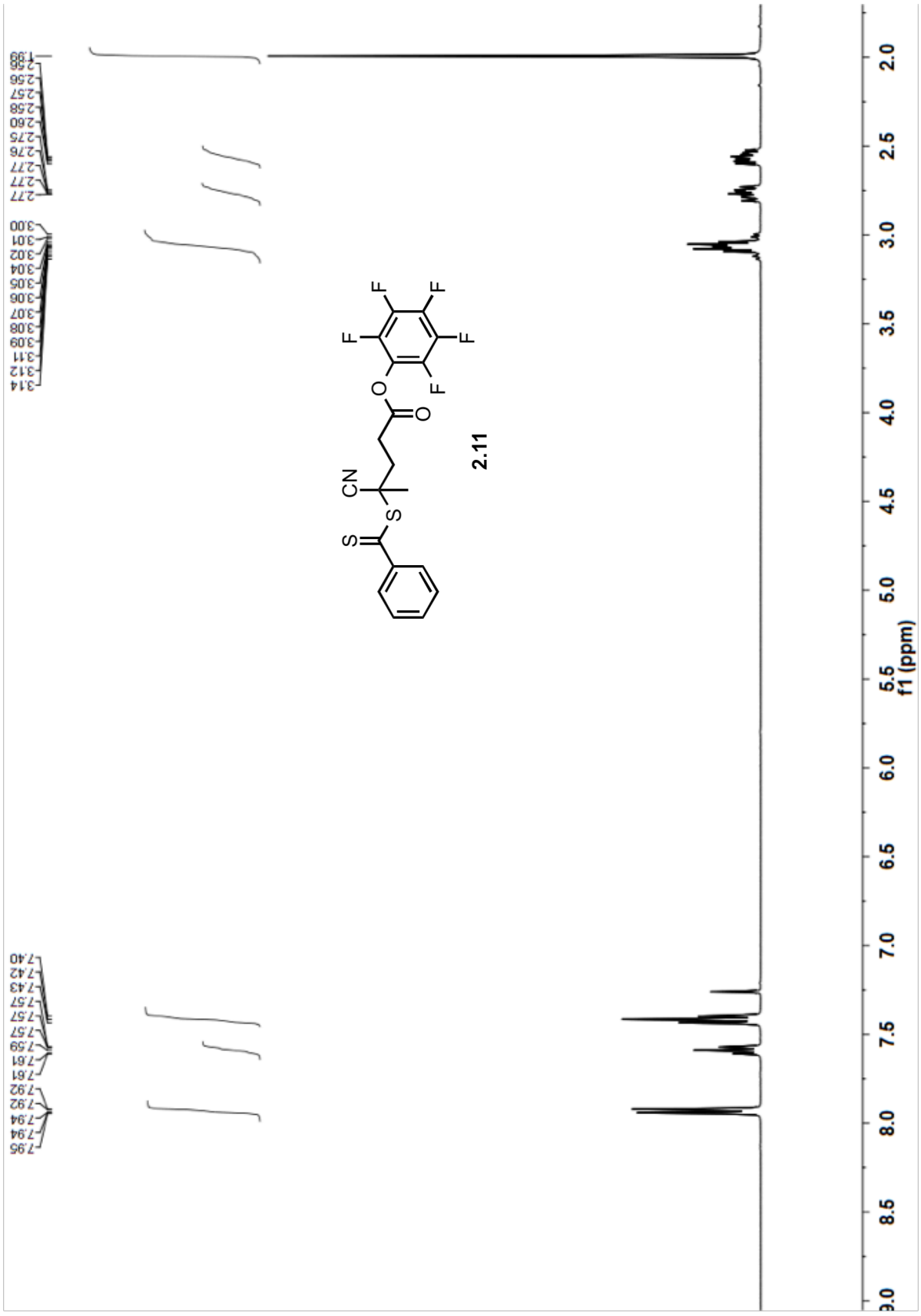

DPPE
O
Br
2.7
f1 (ppm)

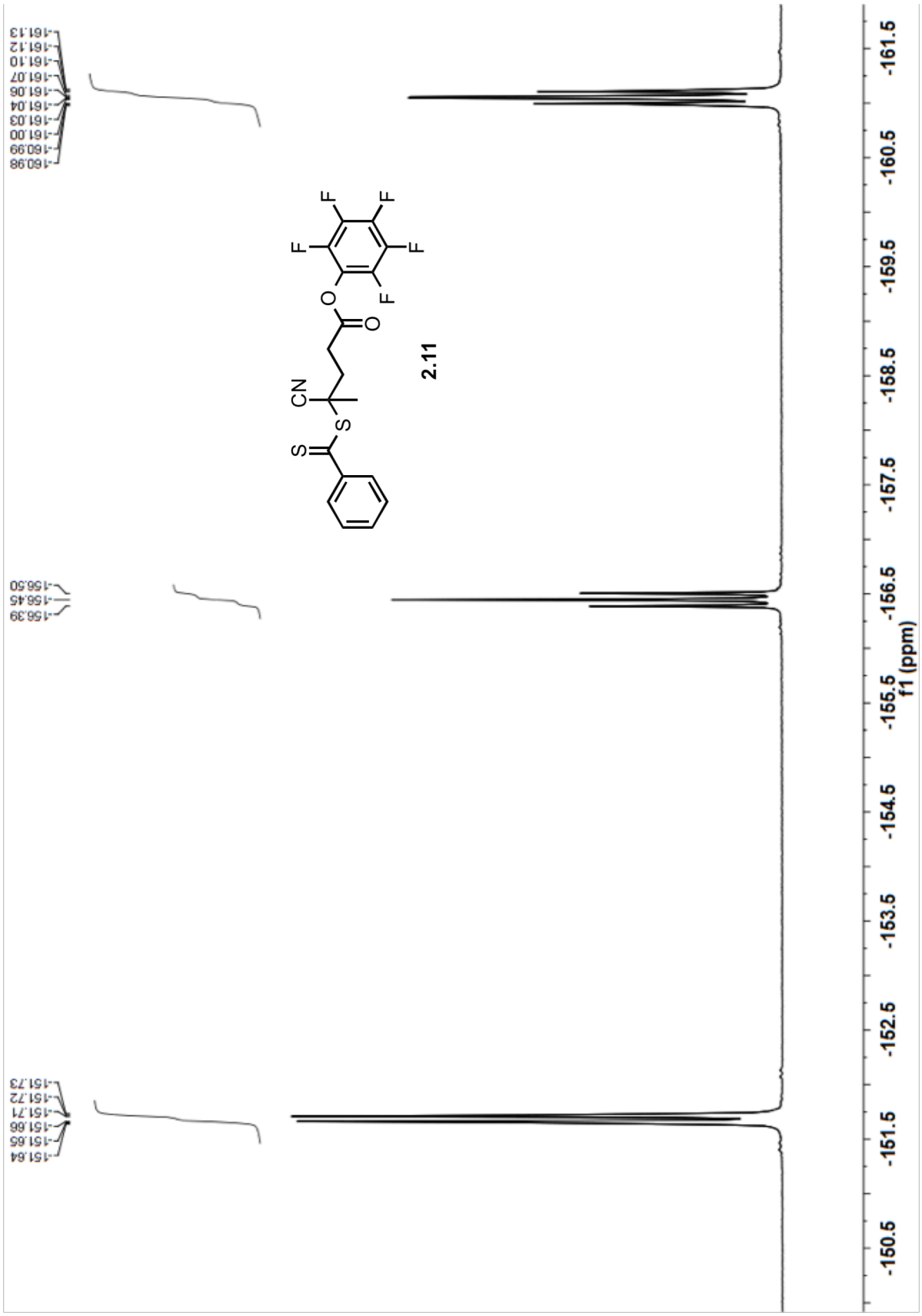

8.69
14.25
22.81
25.00
25.05
29.28
29.49
29.67
29.84
31.19
32.05
32.16
34.23
34.42
45.69
61.37
62.81
63.85
70.46
173.64
173.27
172.32
DPPE
O
Br
2.7
f1 (ppm)

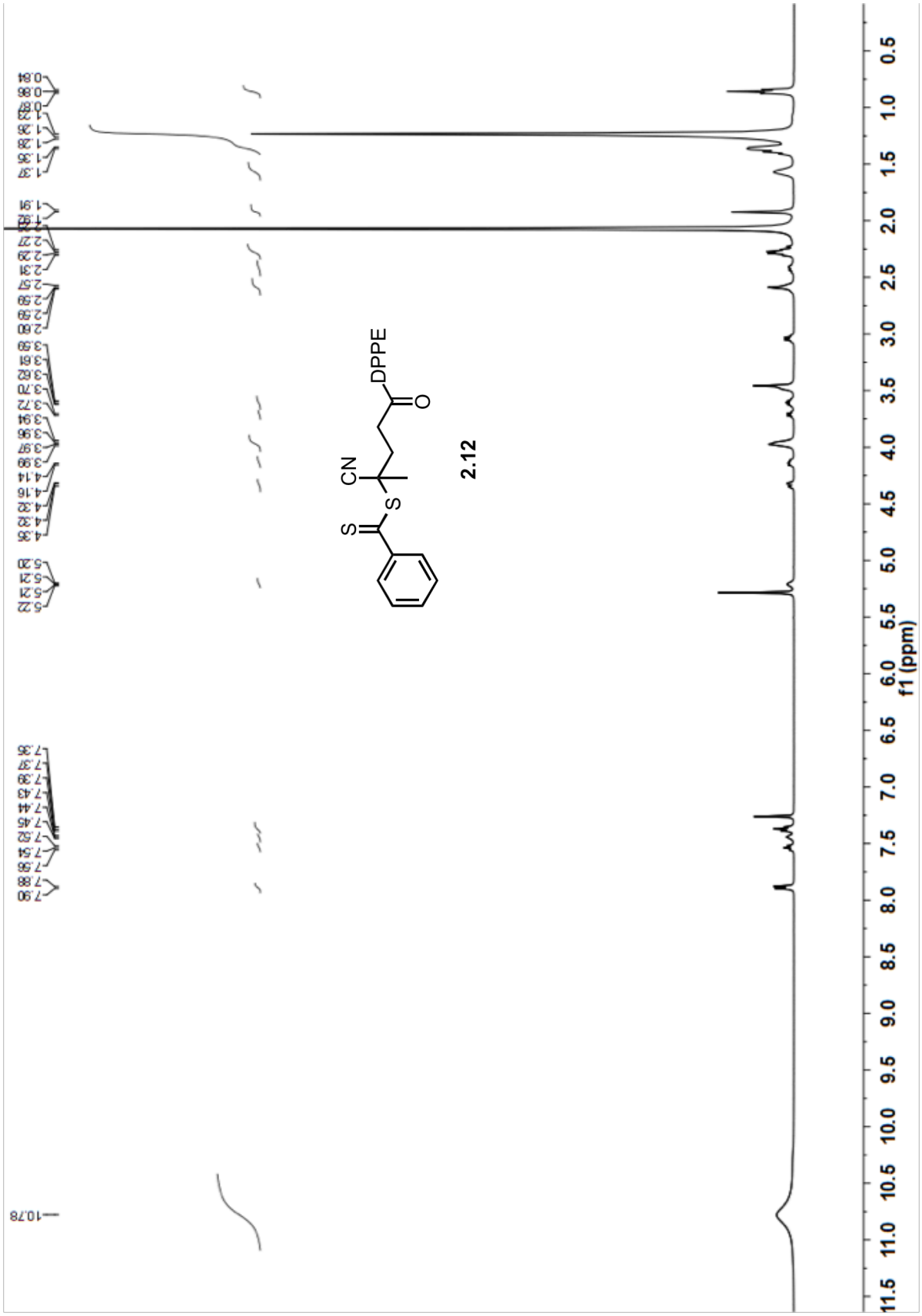

0.40
DPPE
O
Br
2.7
50
45
40
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
f1 (ppm)

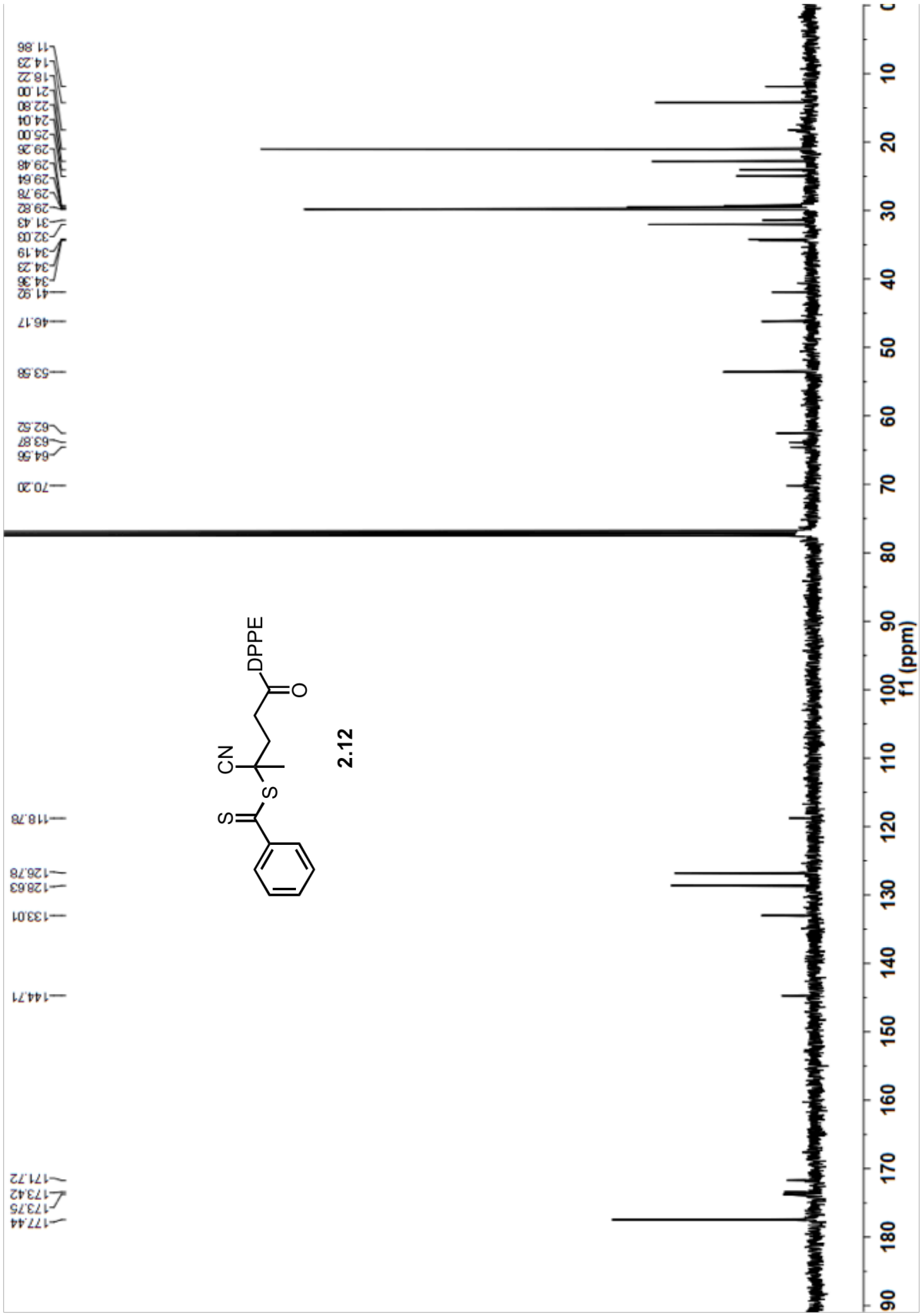

8.10
8.10
8.09
8.08
8.08
8.08
7.64
7.63
7.63
7.62
7.61
7.60
7.60
7.59
7.48
7.47
7.46
7.46
7.44
7.44
S
S
S
S
2.9
8.4
8.3
8.2
8.1
8.0
7.9
7.8
7.7
7.6
7.5
7.4
7.3
7.2
7.1
7.0
6.9
f1 (ppm)

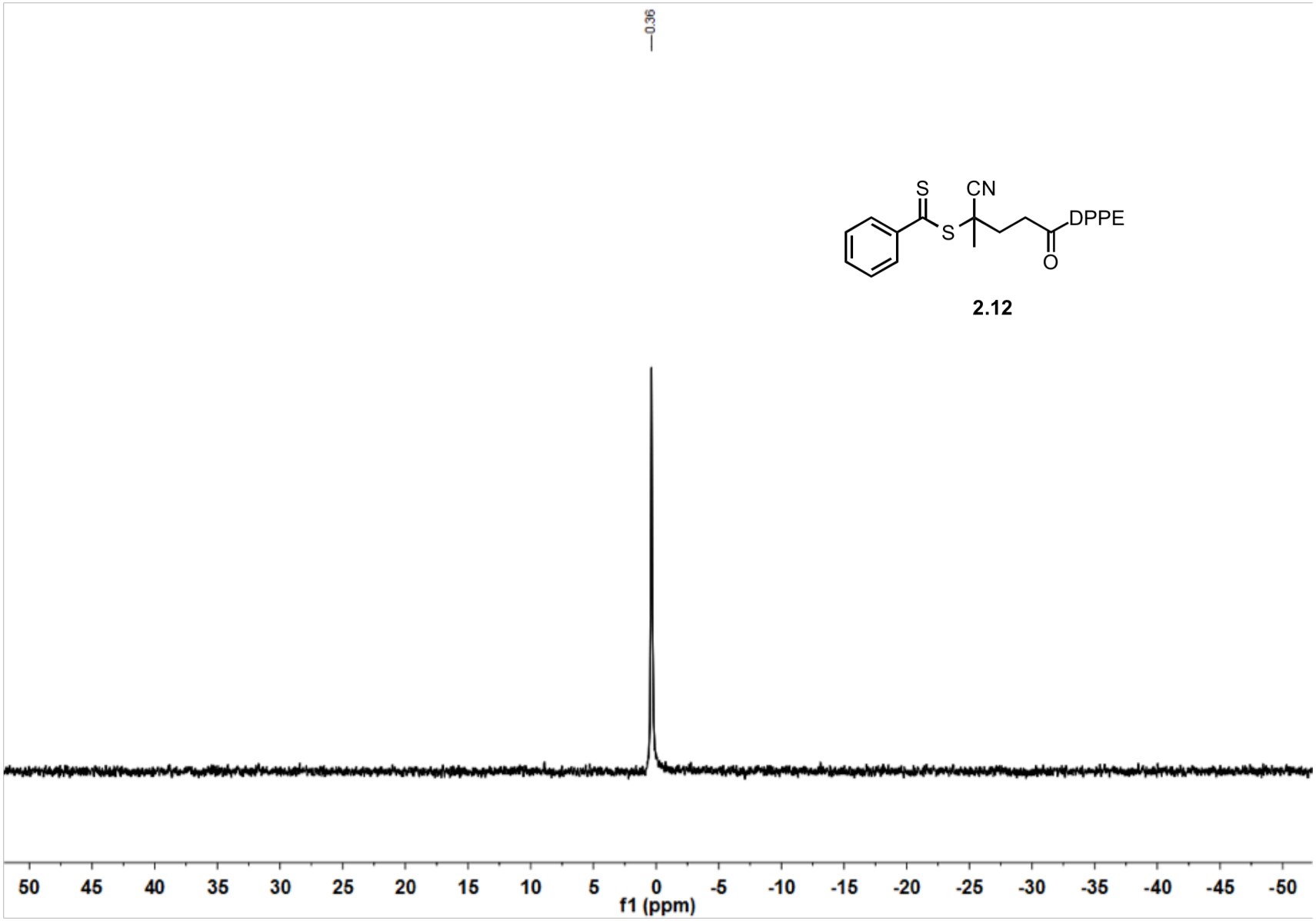

OH
O
CN
S
S
2.10
f1 (ppm)

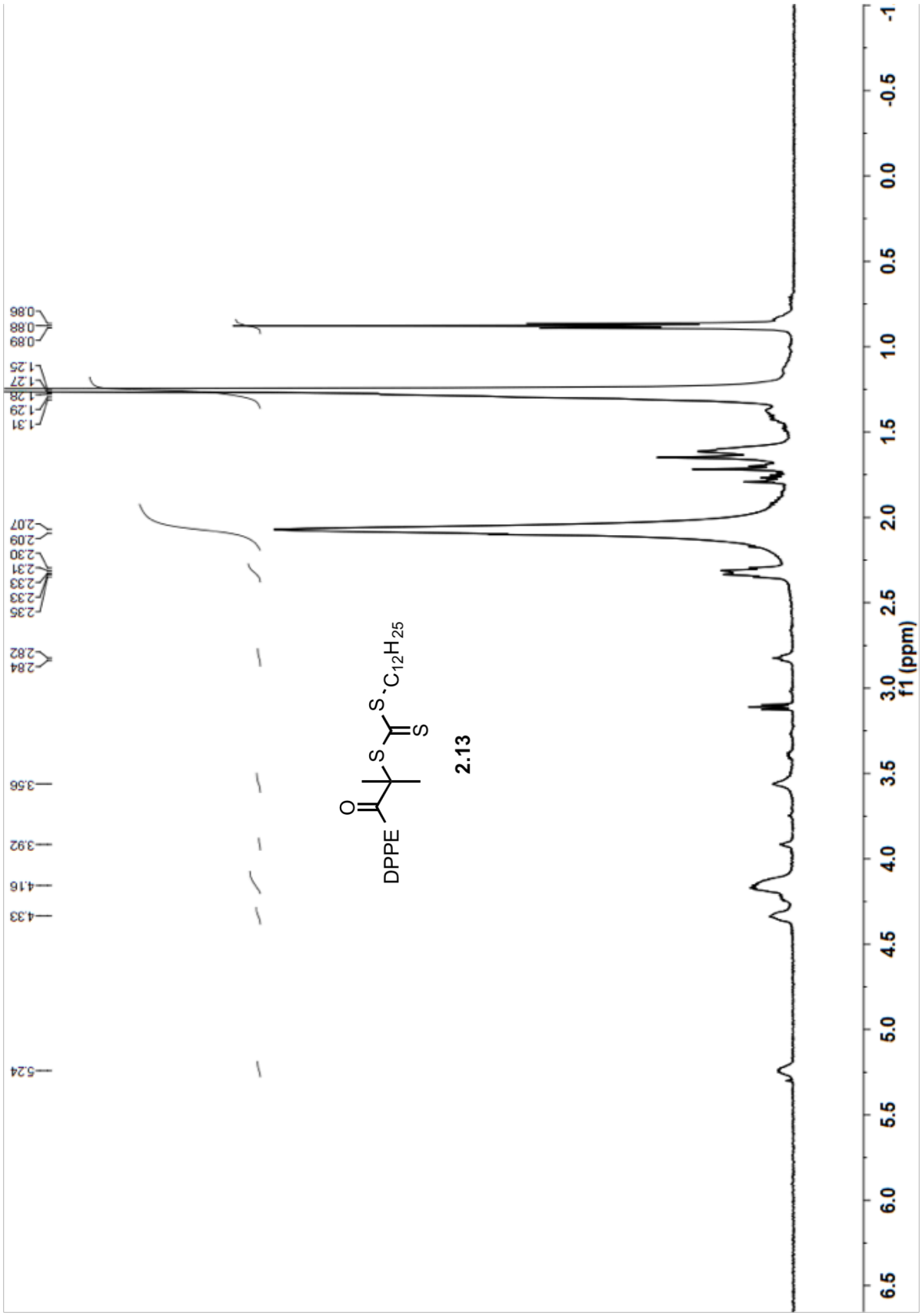

F
F
F
F
F
O
O
CN
S
S
2.11
f1 (ppm)

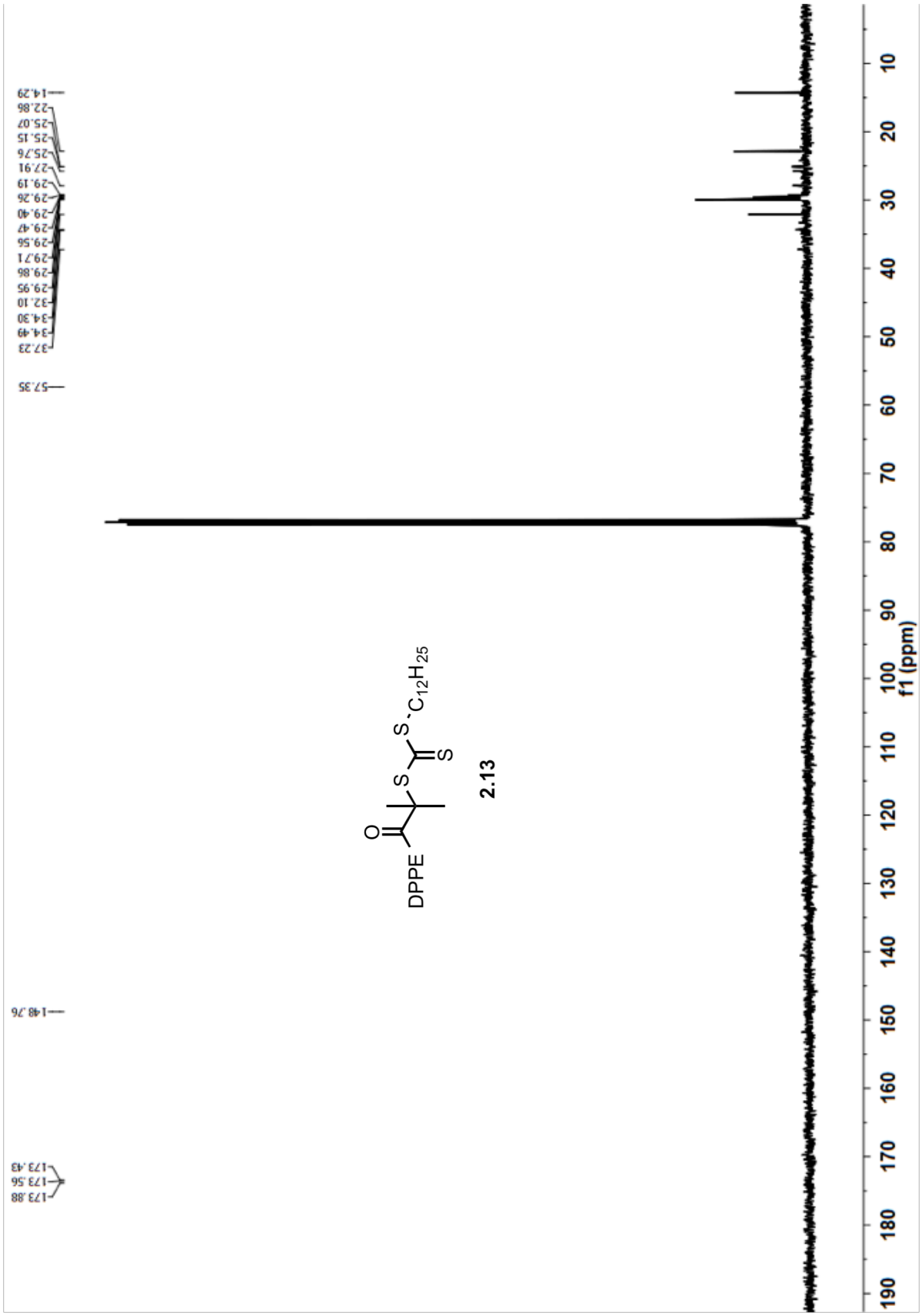

-151.64
-151.65
-151.66
-151.71
-151.72
-151.73
-156.39
-156.45
-156.50
-160.98
-160.99
-161.00
-161.03
-161.04
-161.06
-161.07
-161.10
-161.12
-161.13
F
F
F
F
F
O
O
CN
S
S
2.11
-150.5
-151.5
-152.5
-153.5
-154.5
-155.5
-156.5
-157.5
-158.5
-159.5
-160.5
-161.5
f1 (ppm)

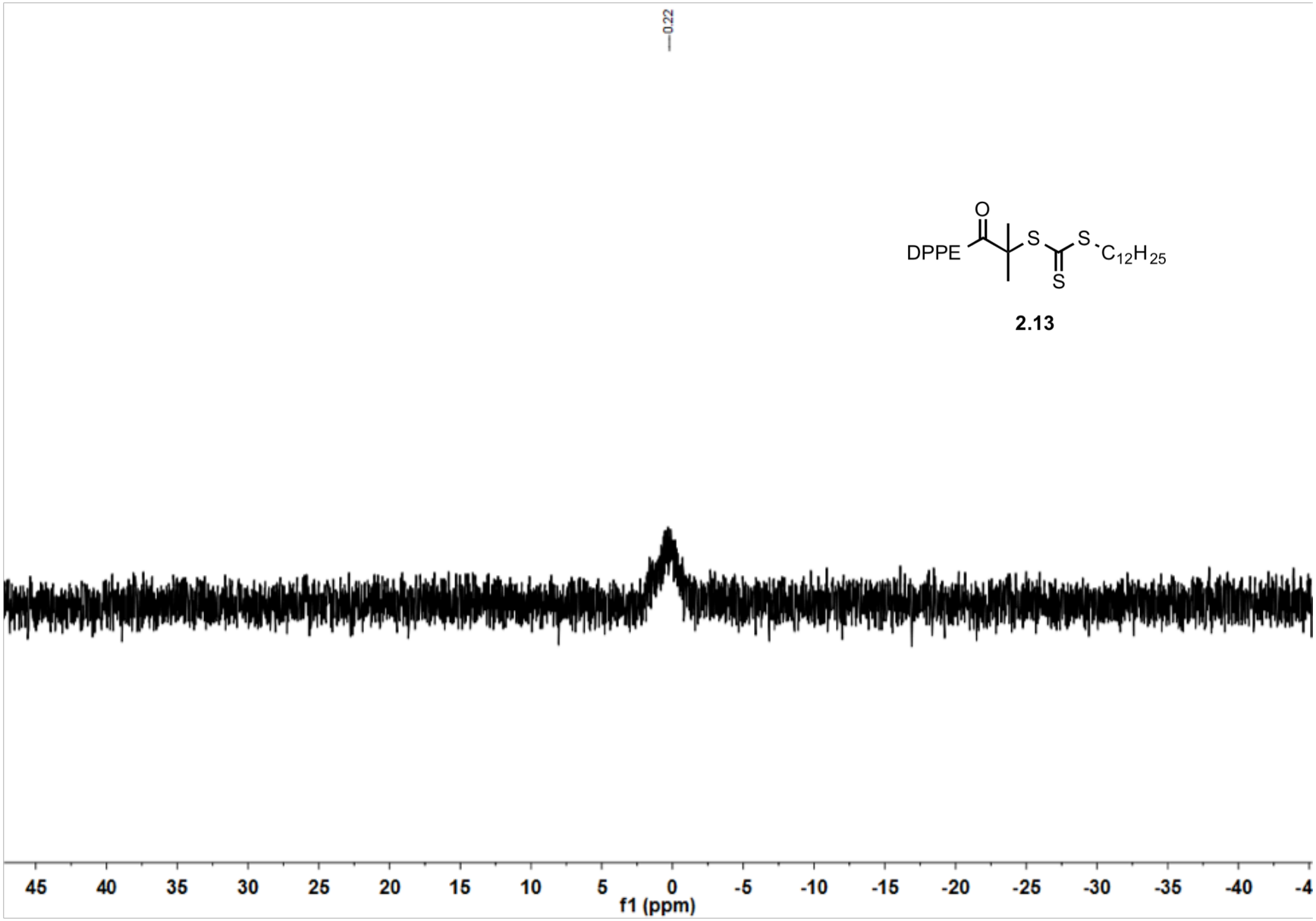

DPPE
CN
S
S
O
2.12
f1 (ppm)


DPPE
CN
S
S
O
2.12
f1 (ppm)

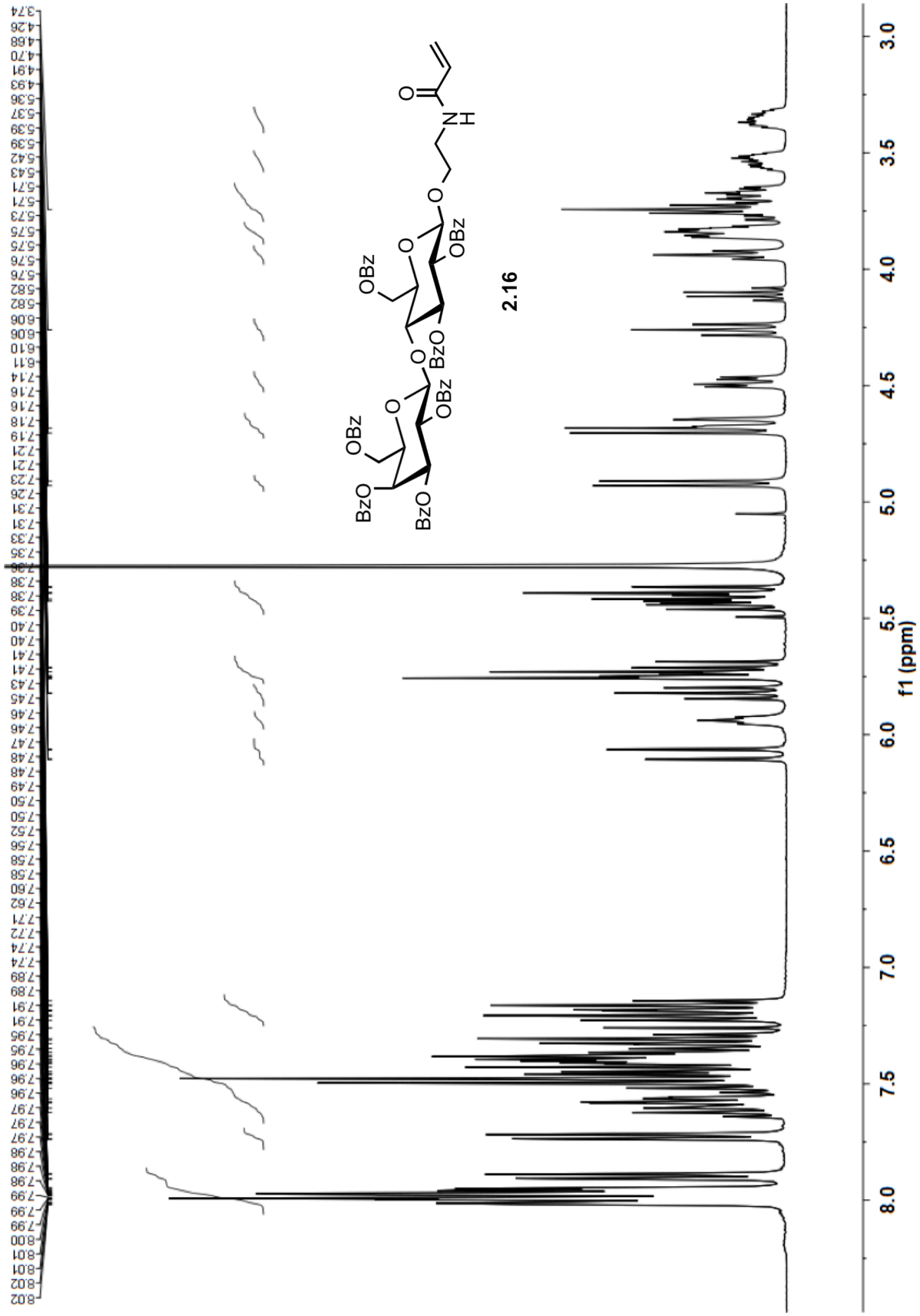

0.36
S
CN
S
DPPE
O
2.12
50 45 40 35 30 25 20 15 10 5 0 -5 -10 -15 -20 -25 -30 -35 -40 -45 -50
f1 (ppm)

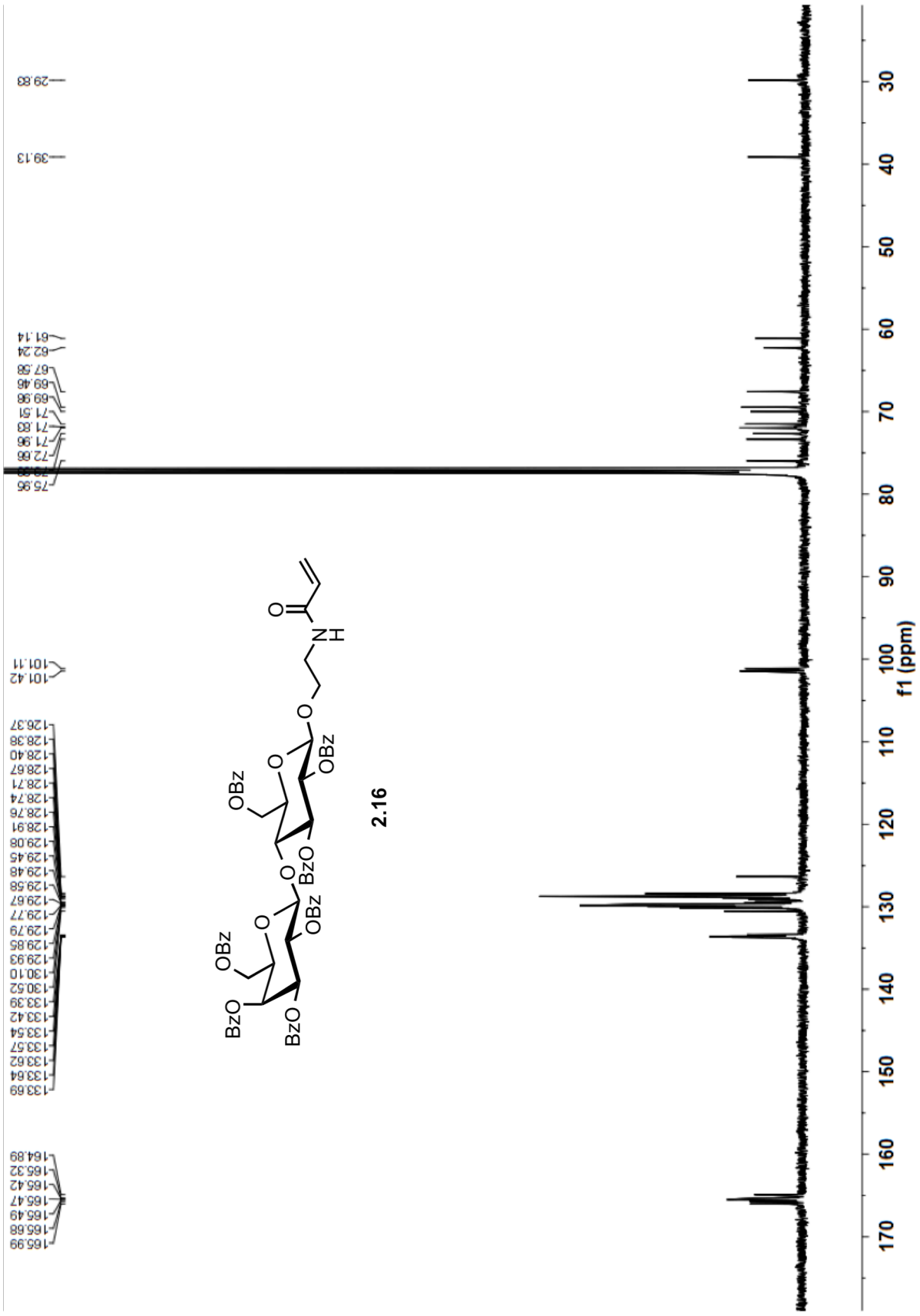

0.86
0.88
0.89
1.25
1.27
1.28
1.29
1.31
2.07
2.09
2.30
2.31
2.33
2.33
2.35
2.82
2.84
3.56
3.92
4.16
4.33
5.24
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1
f1 (ppm)
DPPE
O
S
S
S
$C_{12}H_{25}$
2.13

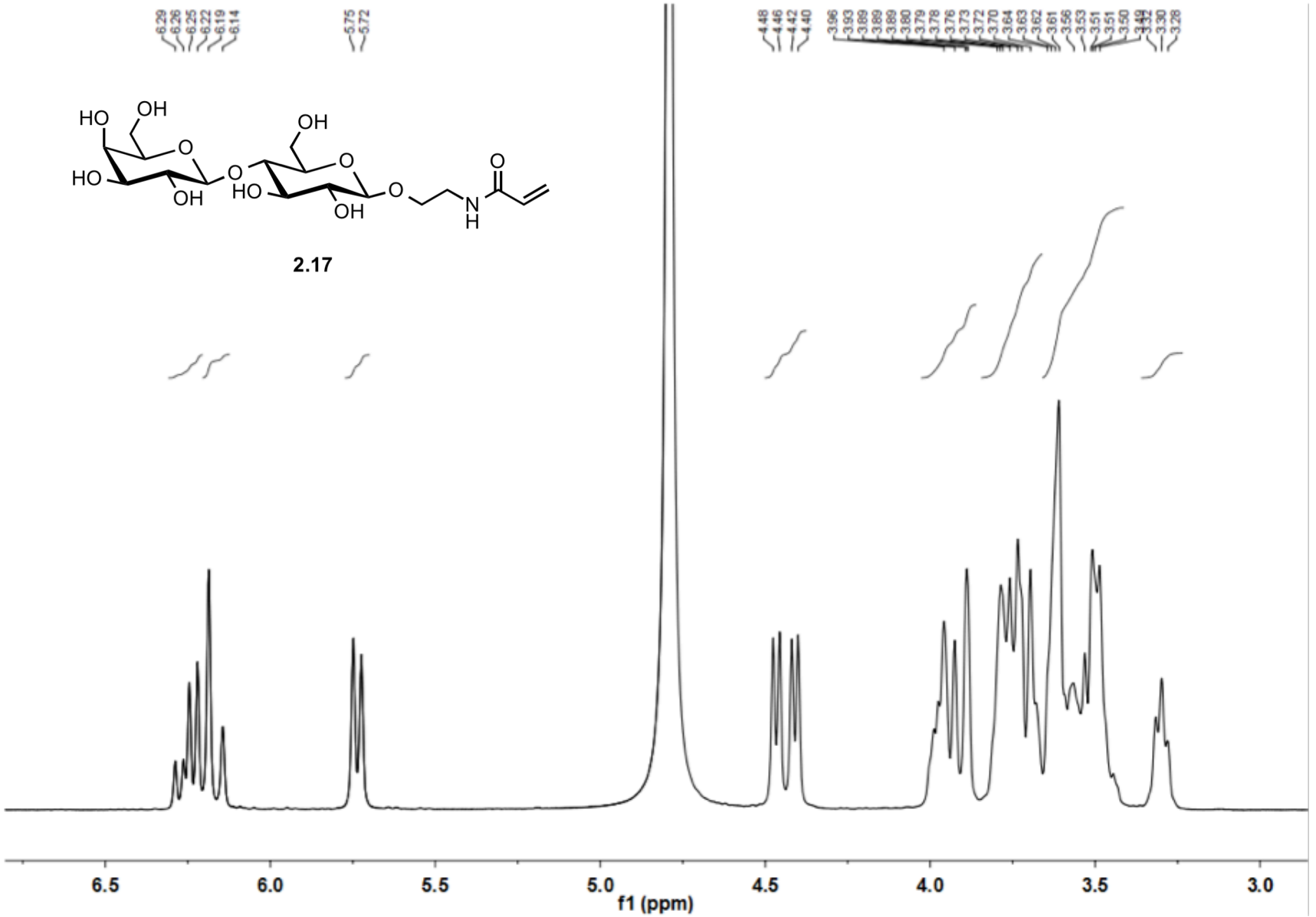

DPPE
O
S
S
S
$C_{12}H_{25}$
2.13
f1 (ppm)

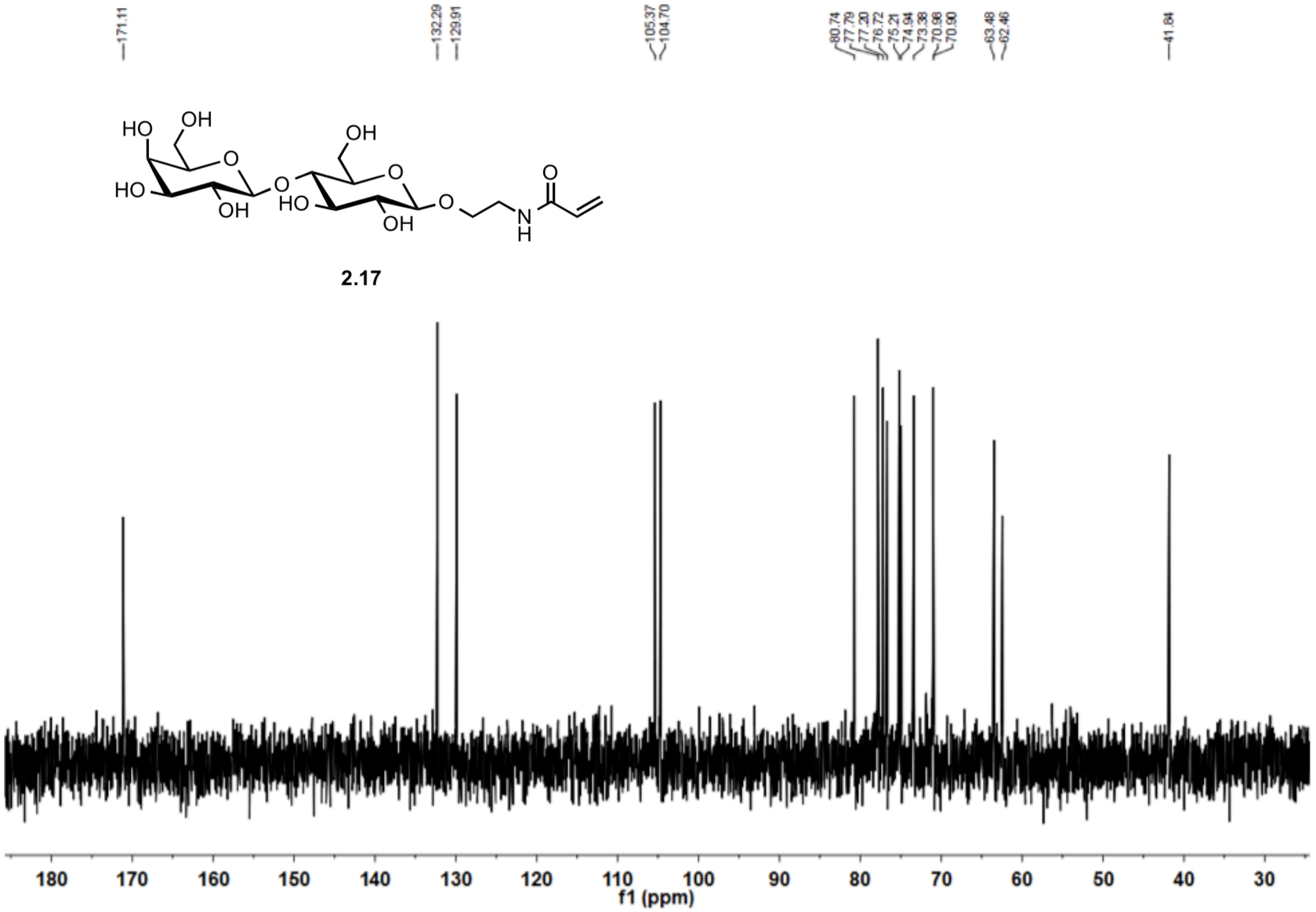

0.22
DPPE
O
S
S
S
$C_{12}H_{25}$
45
40
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40
f1 (ppm)

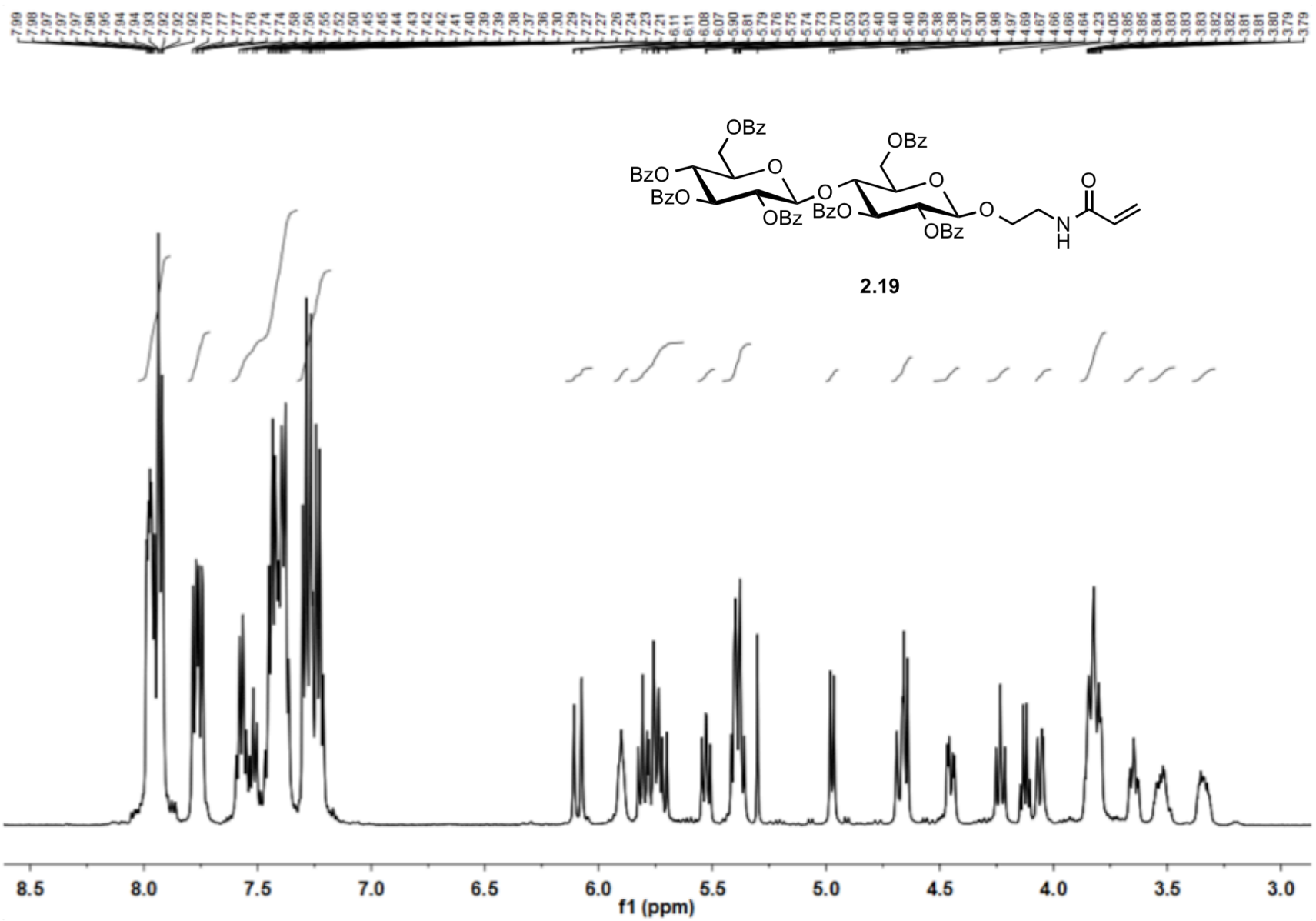
2.13

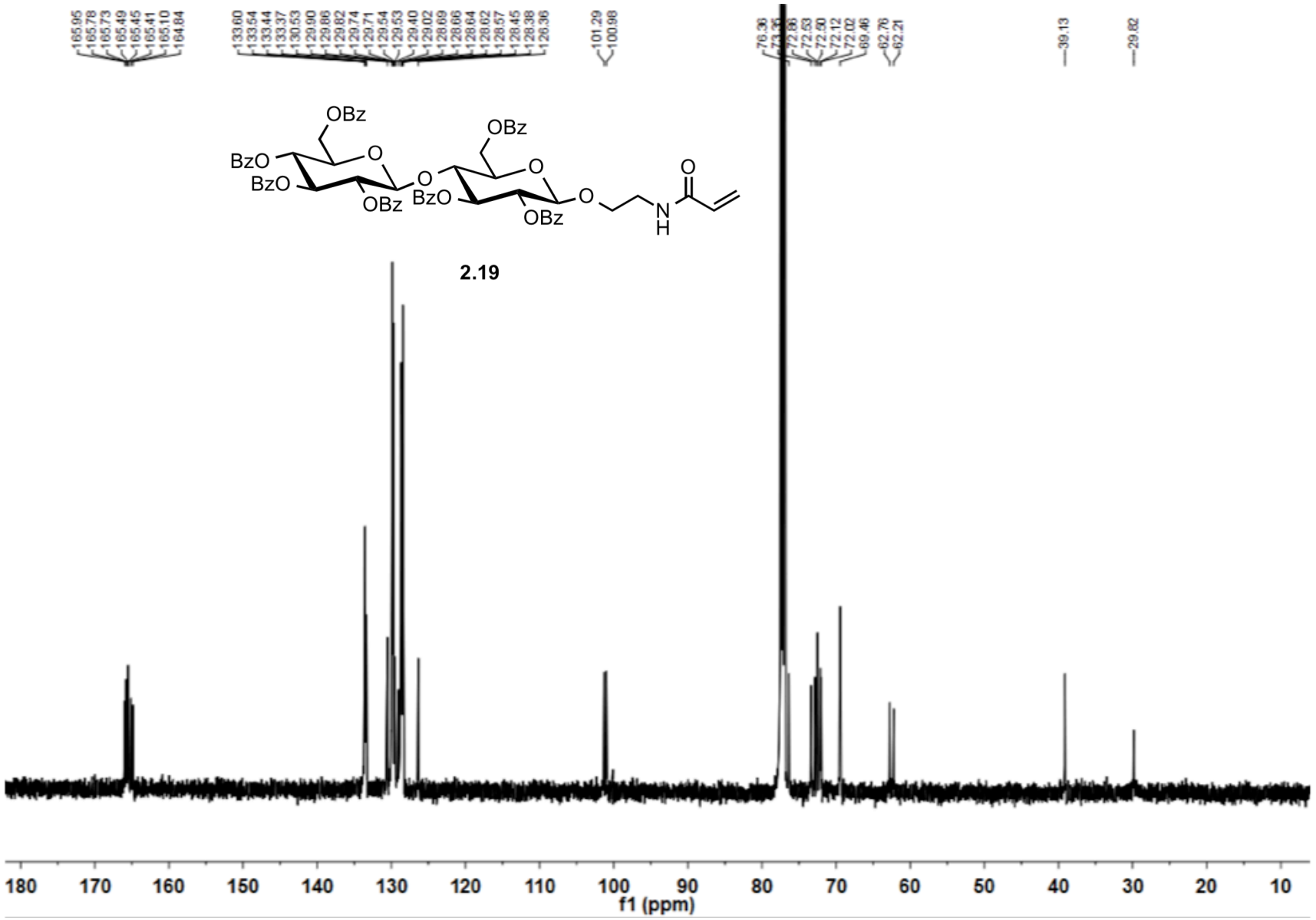

BzO
OBz
O
BzO
OBz
O
BzO
OBz
O
O
H
N
O
2.16
f1 (ppm)

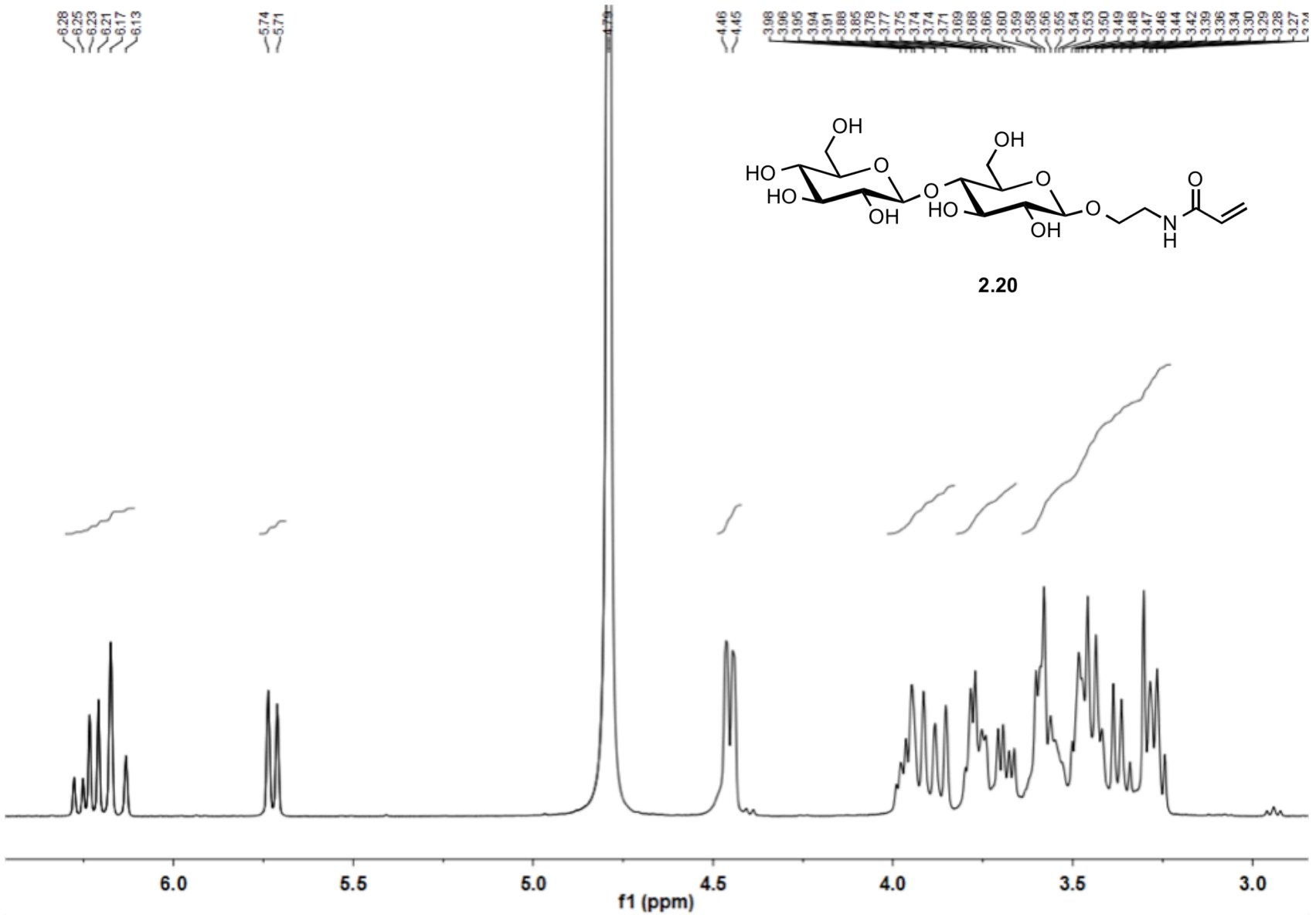

29.83
39.13
61.14
62.24
67.58
69.46
69.96
71.51
71.83
71.96
72.66
75.86
101.11
101.42
126.37
128.38
128.40
128.67
128.71
128.74
128.76
128.91
129.08
129.45
129.48
129.58
129.67
129.77
129.79
129.85
129.93
130.10
130.52
133.39
133.42
133.54
133.57
133.62
133.64
133.69
164.89
165.32
165.42
165.47
165.49
165.68
165.99
OBz
BzO
OBz
BzO
OBz
BzO
OBz
O
N
H
2.16
f1 (ppm)

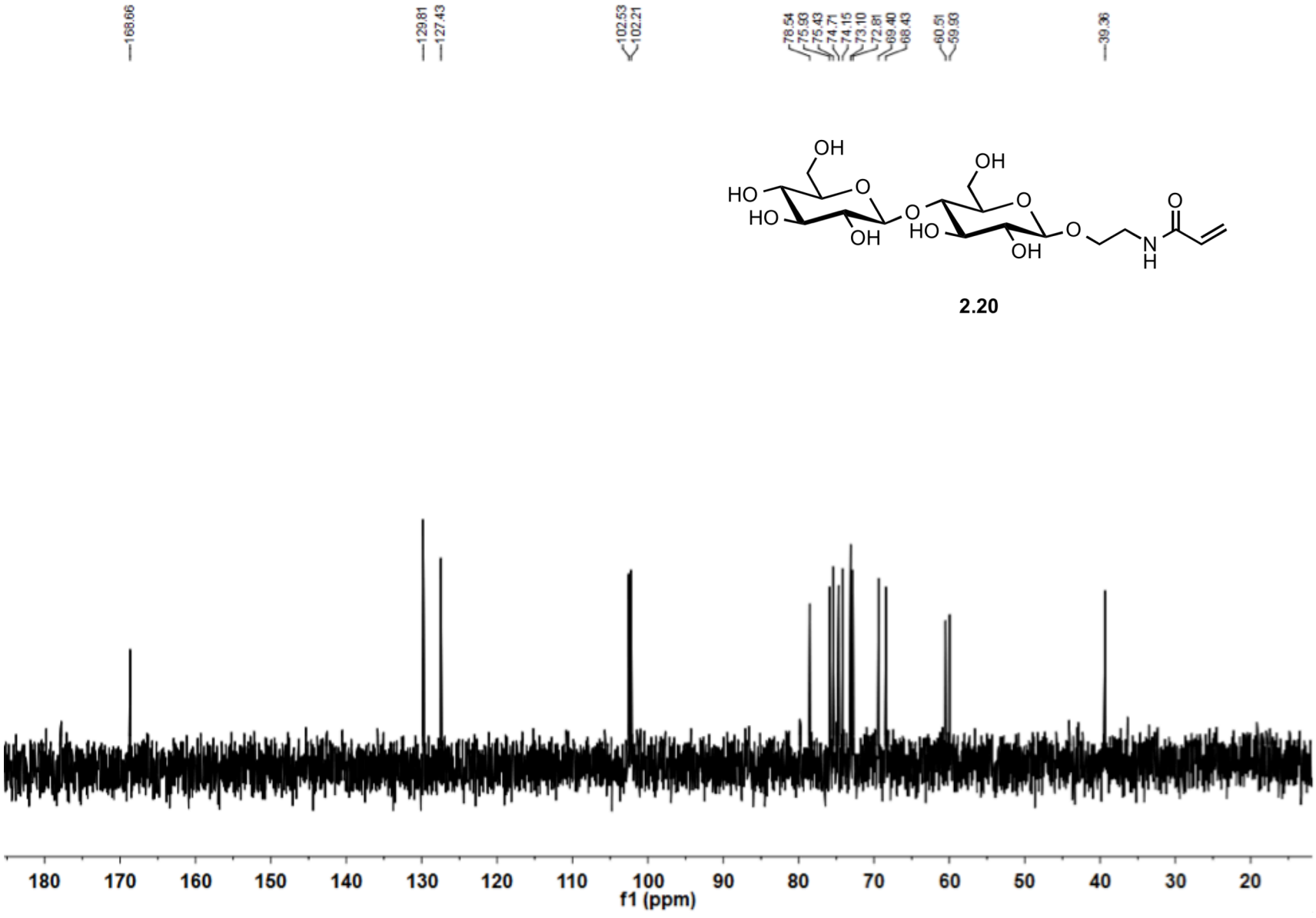

6.29
6.26
6.25
6.22
6.19
6.14
5.75
5.72
4.48
4.46
4.42
4.40
3.96
3.93
3.89
3.89
3.89
3.80
3.79
3.78
3.76
3.73
3.72
3.70
3.64
3.63
3.62
3.61
3.56
3.53
3.51
3.51
3.50
3.32
3.30
3.28
HO
OH
O
HO
OH
O
OH
HO
O
OH
O
N
H
O
2.17
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
f1 (ppm)

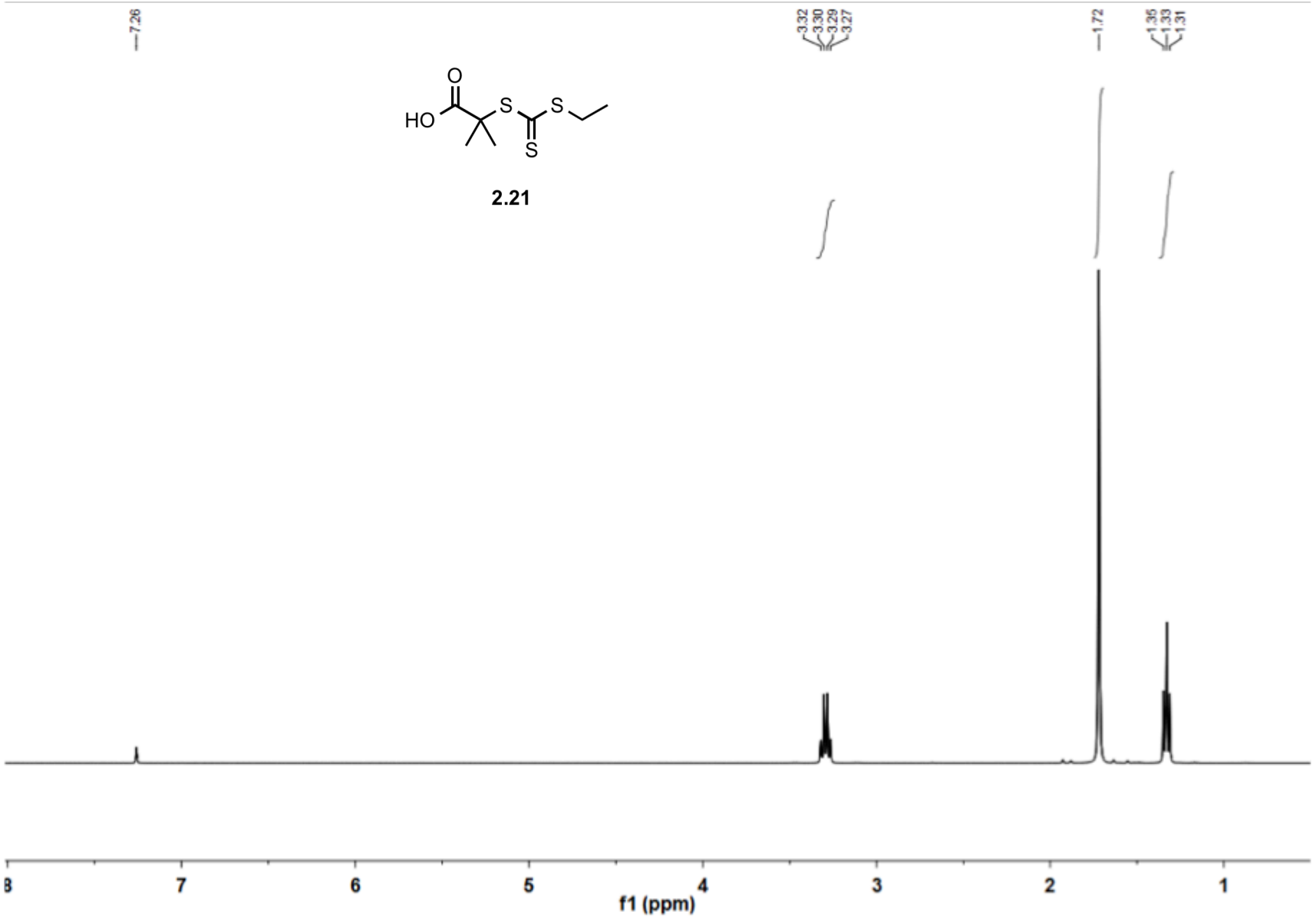

171.11
132.29
129.91
105.37
104.70
80.74
77.79
77.20
76.72
75.21
74.94
73.38
70.98
70.90
63.48
62.46
41.84
HO
OH
O
HO
OH
O
HO
OH
O
OH
O
N
H
O
2.17
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
f1 (ppm)

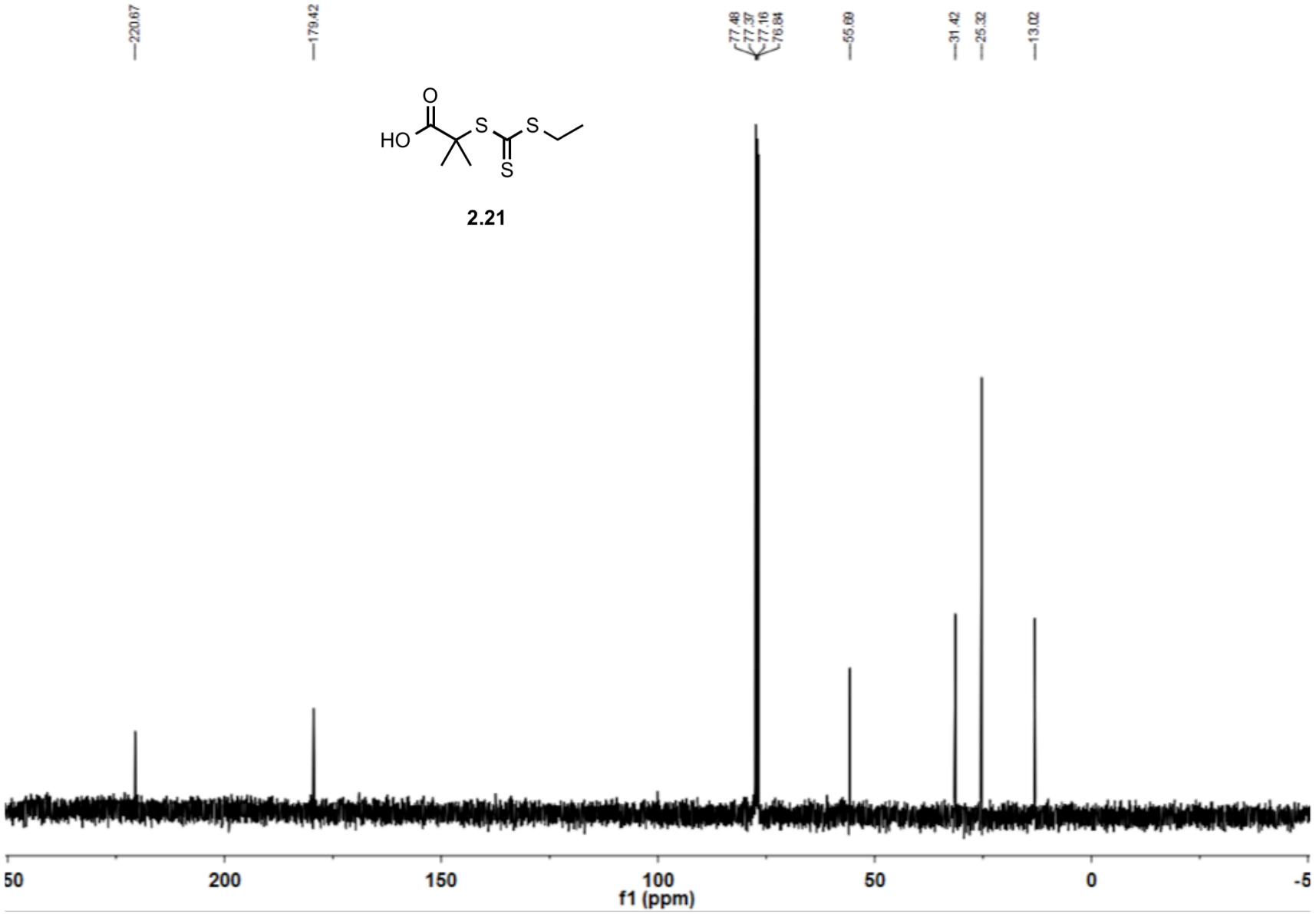

OBz
BzO
BzO
OBz
O
O
OBz
BzO
O
OBz
O
N
H
O
2.19
f1 (ppm)
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0

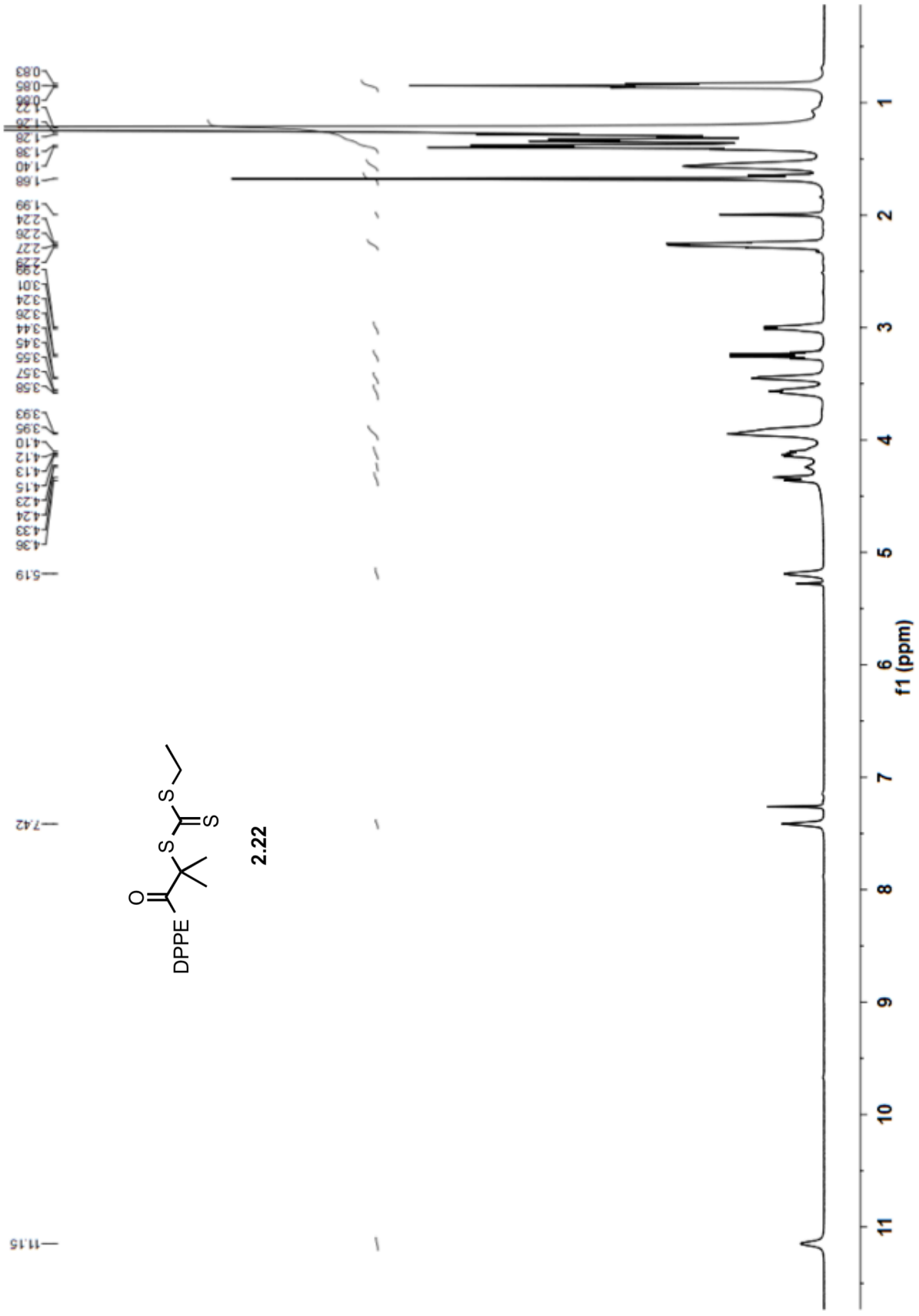

165.95
165.78
165.73
165.49
165.45
165.41
165.10
164.84
133.60
133.54
133.44
133.37
130.53
129.90
129.86
129.82
129.74
129.71
129.54
129.53
129.40
129.02
128.69
128.66
128.64
128.62
128.57
128.45
128.38
126.36
101.29
100.98
76.30
73.35
72.88
72.53
72.50
72.12
72.02
69.46
62.76
62.21
39.13
29.62
OBz
BzO
BzO
O
OBz
O
OBz
BzO
O
OBz
O
N
H
O
2.19
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)

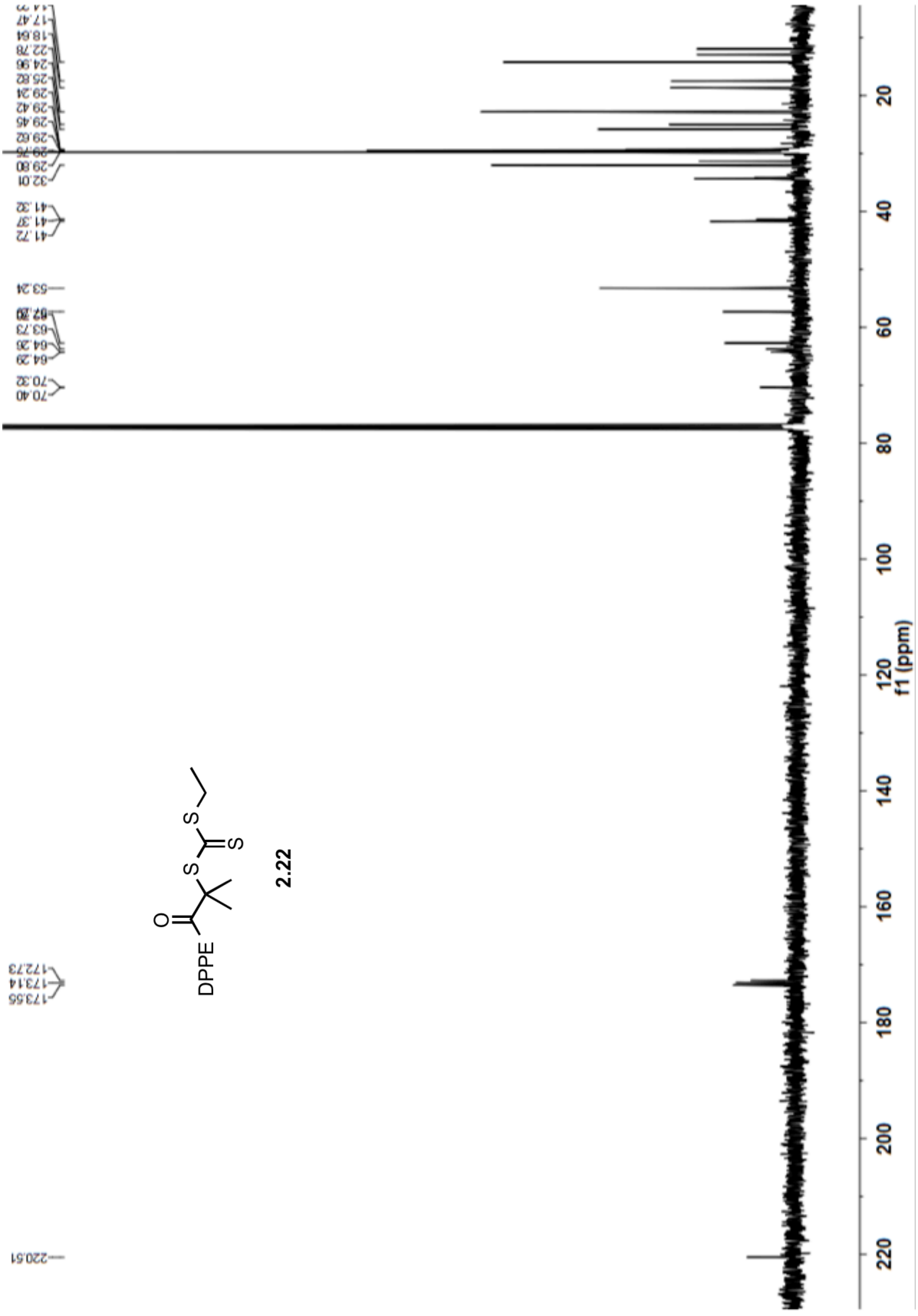

6.28
6.25
6.23
6.21
6.17
6.13
5.74
5.71
4.79
4.46
4.45
3.98
3.96
3.95
3.94
3.91
3.88
3.85
3.78
3.77
3.75
3.74
3.74
3.71
3.69
3.68
3.66
3.60
3.59
3.58
3.56
3.55
3.54
3.53
3.50
3.49
3.48
3.47
3.46
3.44
3.42
3.39
3.36
3.34
3.30
3.29
3.28
3.27
OH
HO
HO
OH
O
O
OH
HO
OH
O
O
N
H
O
2.20
6.0
5.5
5.0
4.5
4.0
3.5
3.0
f1 (ppm)

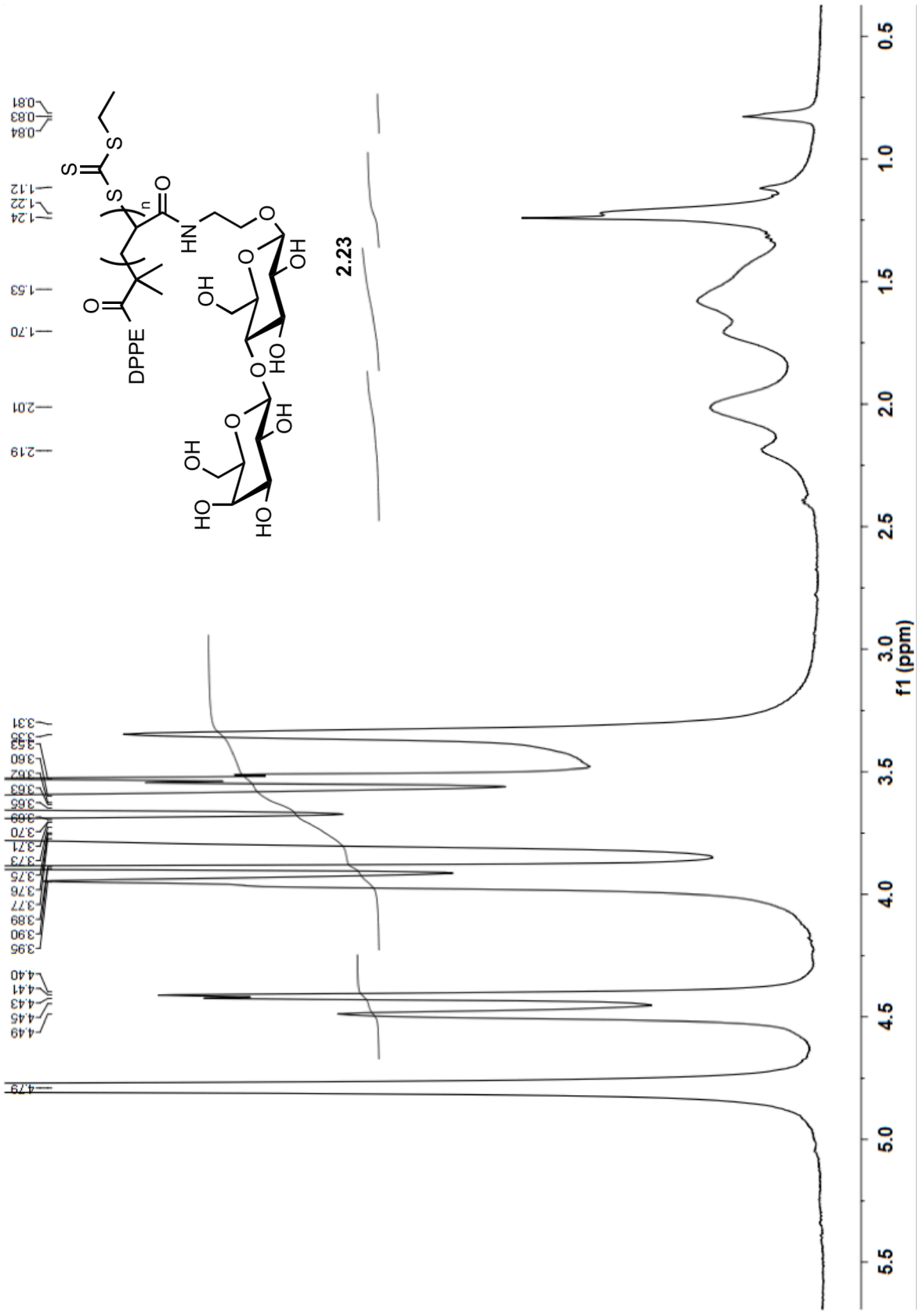

168.66
129.81
127.43
102.53
102.21
78.54
75.93
75.43
74.71
74.15
73.10
72.81
69.40
68.43
60.51
59.93
39.36
OH
HO
HO
OH
O
O
OH
HO
OH
O
O
N
H
O
2.20
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
f1 (ppm)

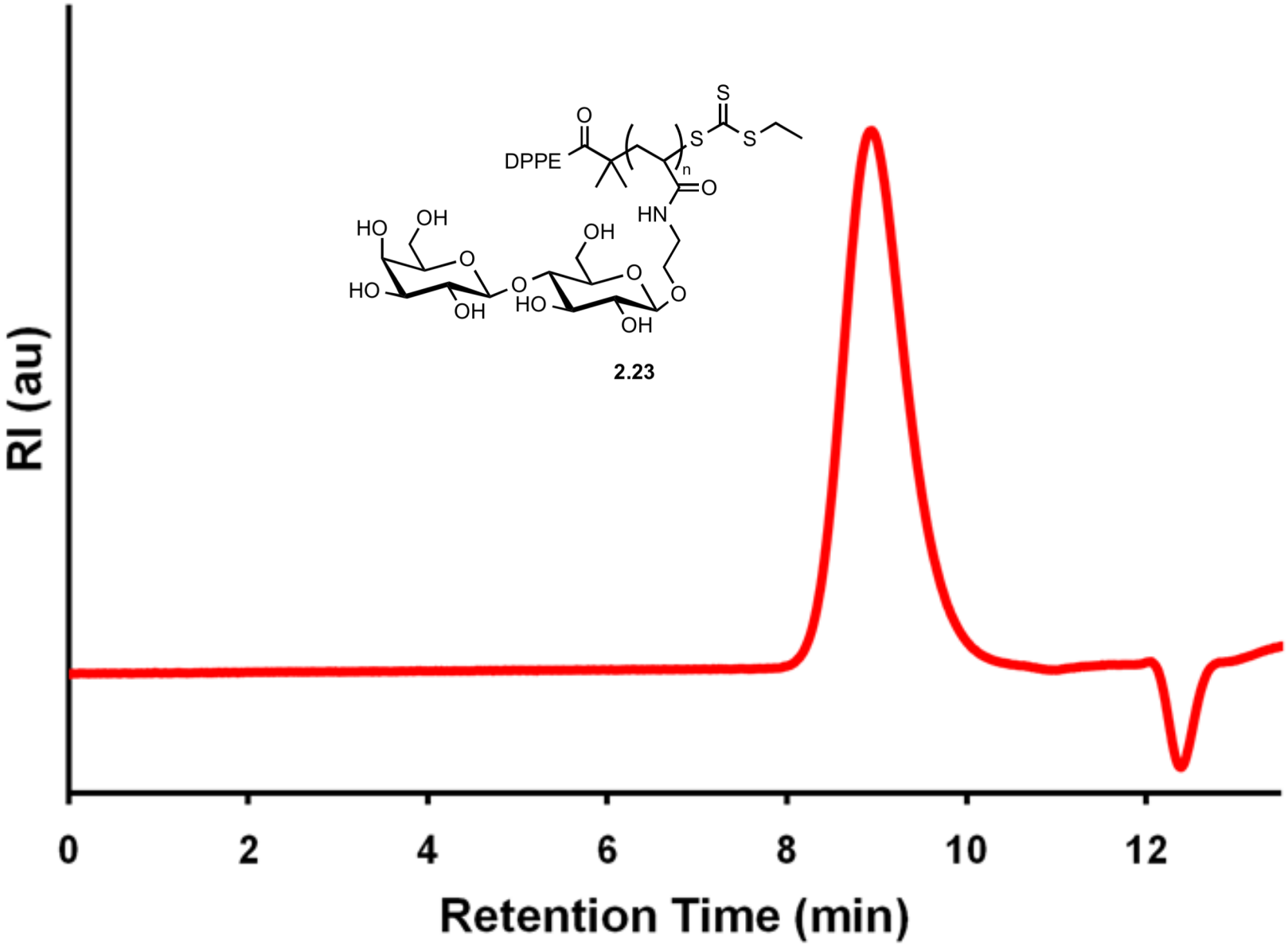

7.26
3.32
3.30
3.29
3.27
1.72
1.35
1.33
1.31
HO
O
S
S
S
2.21
8
7
6
5
4
3
2
1
f1 (ppm)


220.67
179.42
77.48
77.37
77.16
76.84
55.69
31.42
25.32
13.02
HO
O
S
S
S
2.21
50
200
150
100
50
0
-5
f1 (ppm)

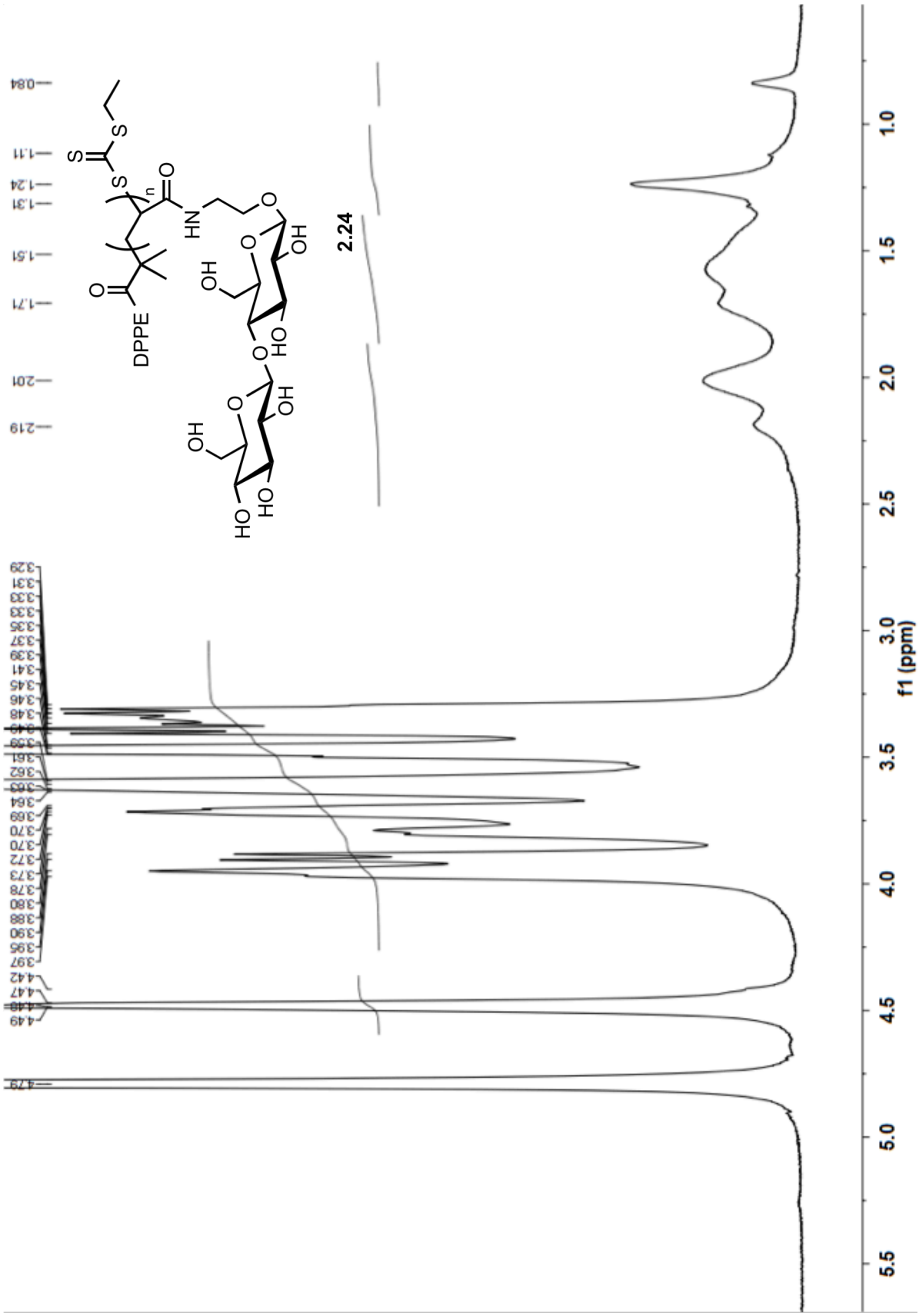

DPPE
O
S
S
S
2.22
11.15
7.42
5.19
f1 (ppm)

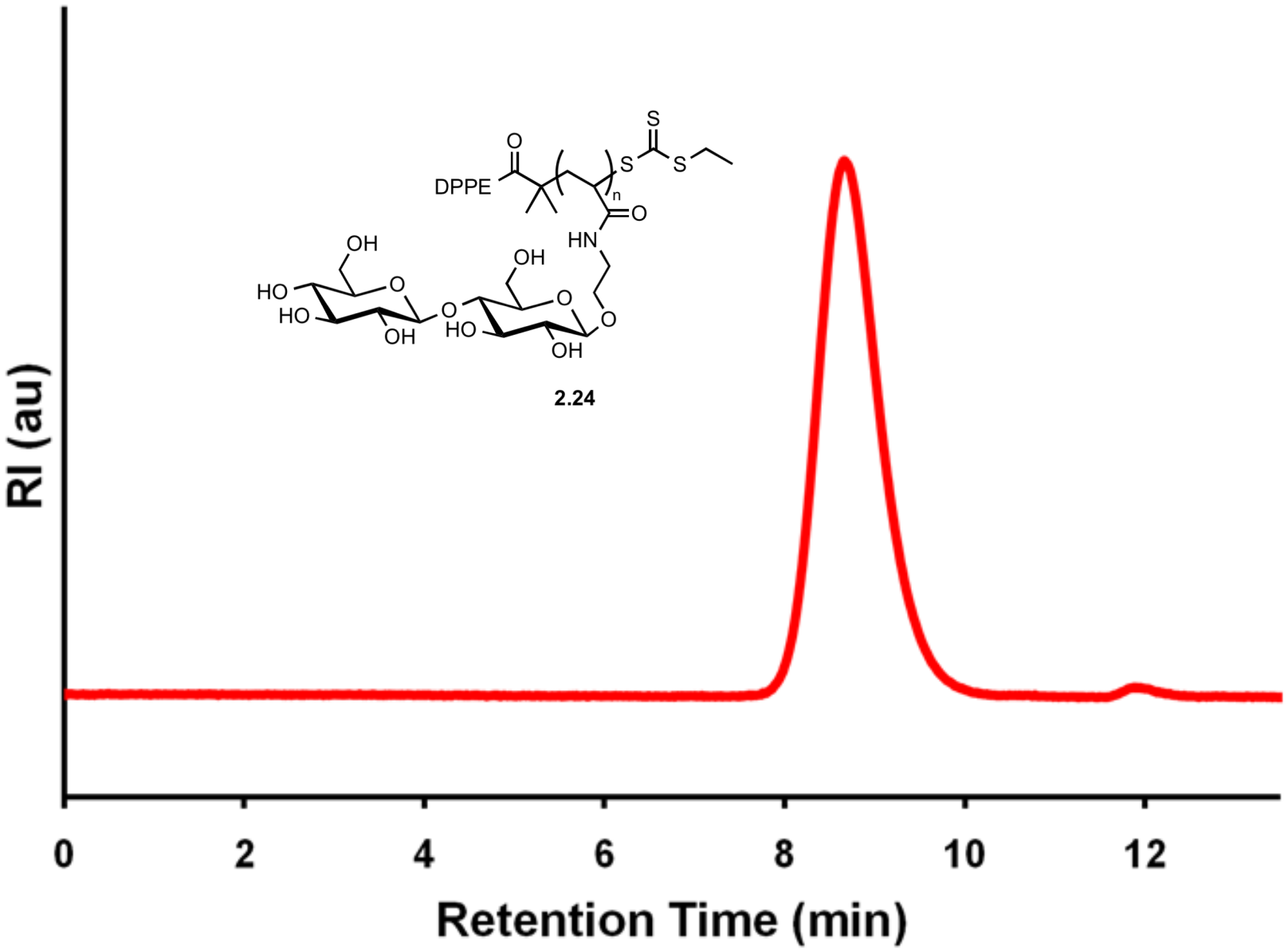

DPPE
O
S
S
S
2.22
220.51
173.55
173.14
172.73
70.40
70.32
64.29
64.26
63.73
53.24
41.72
41.37
41.32
32.01
29.80
29.76
29.62
29.45
29.42
29.24
25.82
24.96
22.78
18.64
17.47
f1 (ppm)


4.79
4.49
4.45
4.43
4.41
4.40
3.95
3.90
3.89
3.77
3.76
3.75
3.73
3.71
3.70
3.69
3.65
3.63
3.62
3.60
3.59
3.36
3.31
2.19
2.01
1.70
1.53
1.24
1.22
1.12
0.84
0.83
0.81
2.23
DPPE
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

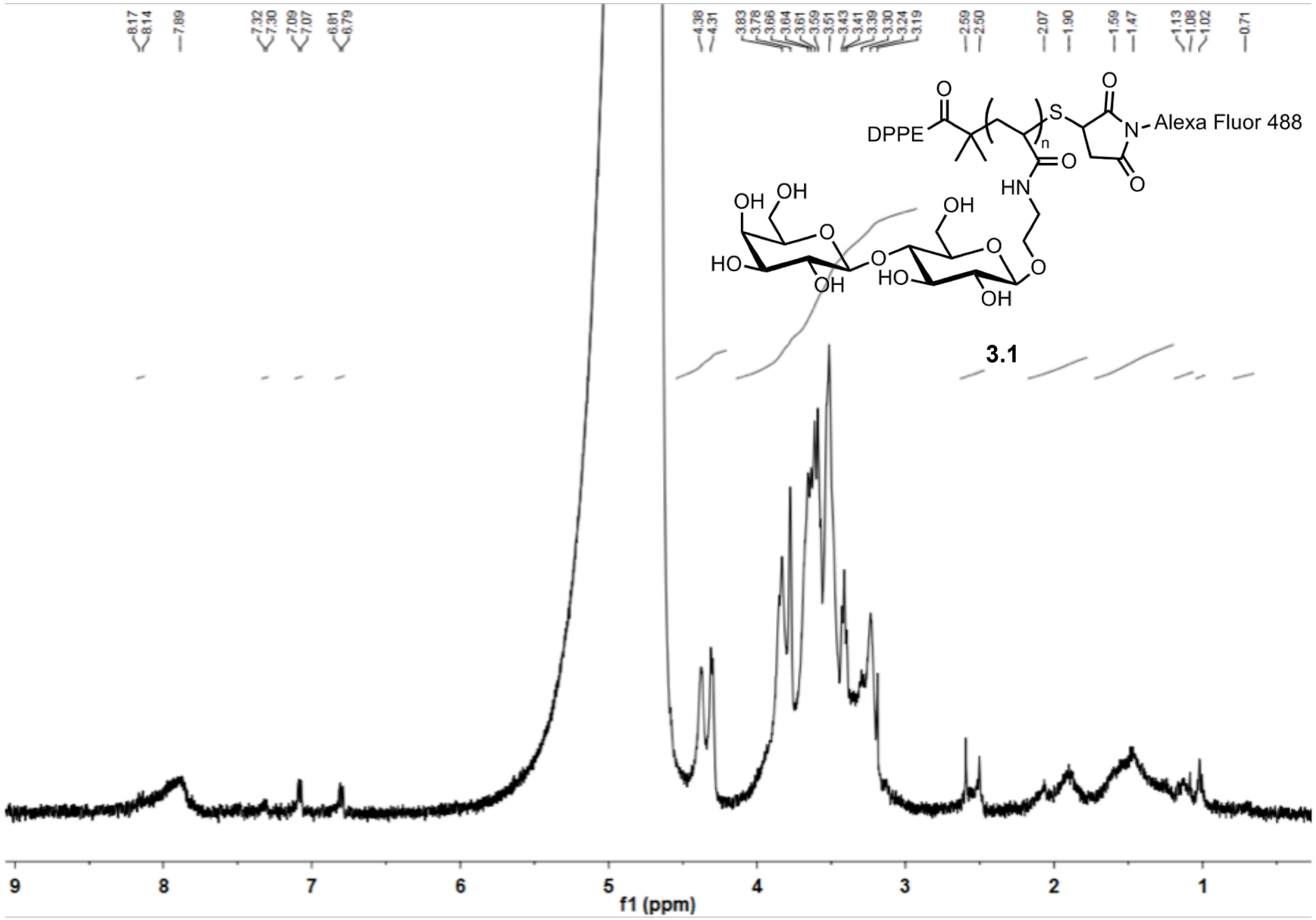

RI (au)
0
2
4
6
8
10
12
Retention Time (min)
DPPE
O
S
S
S
n
O
HN
HO
OH
O
HO
OH
O
OH
HO
O
O
OH

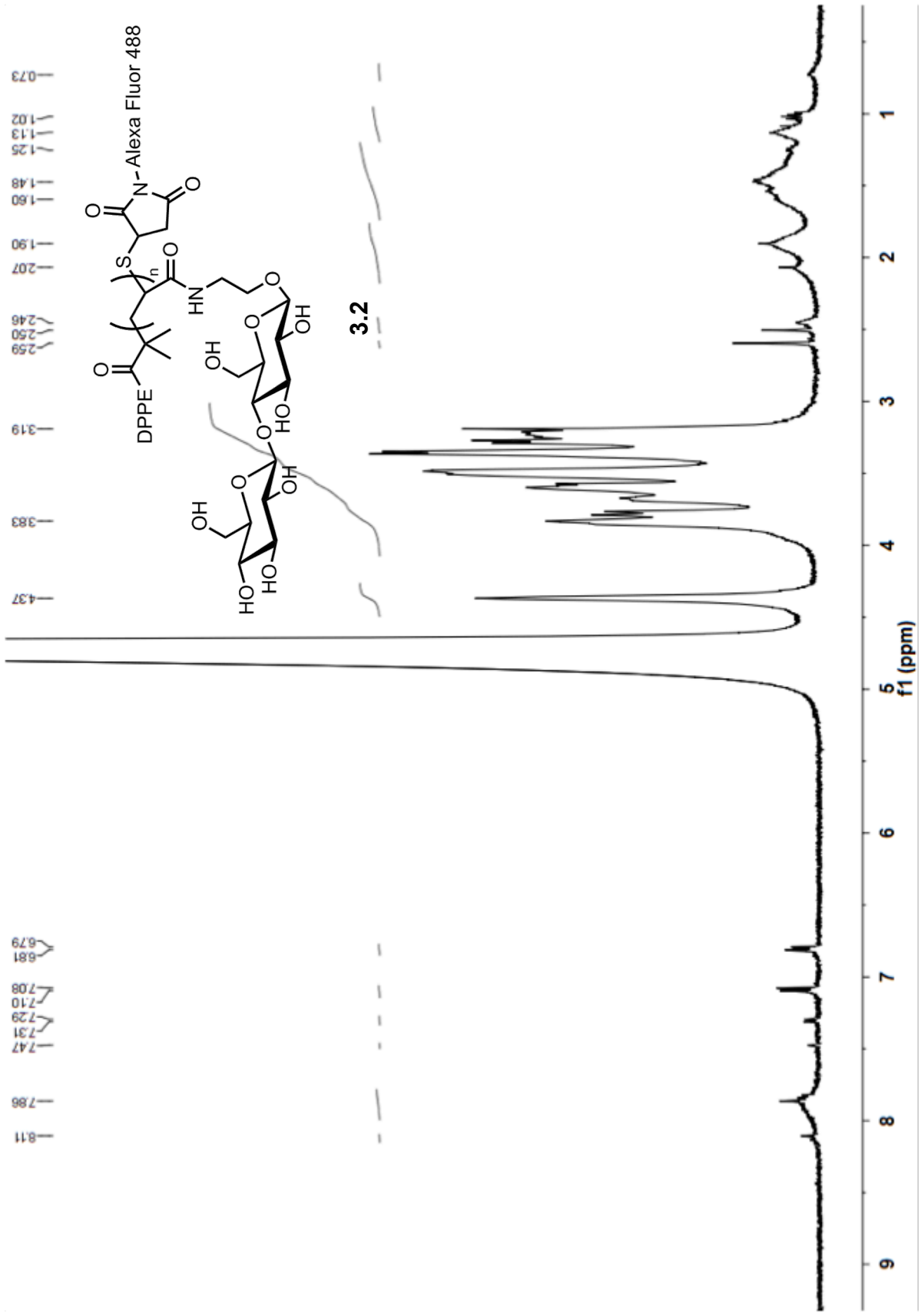
2.23

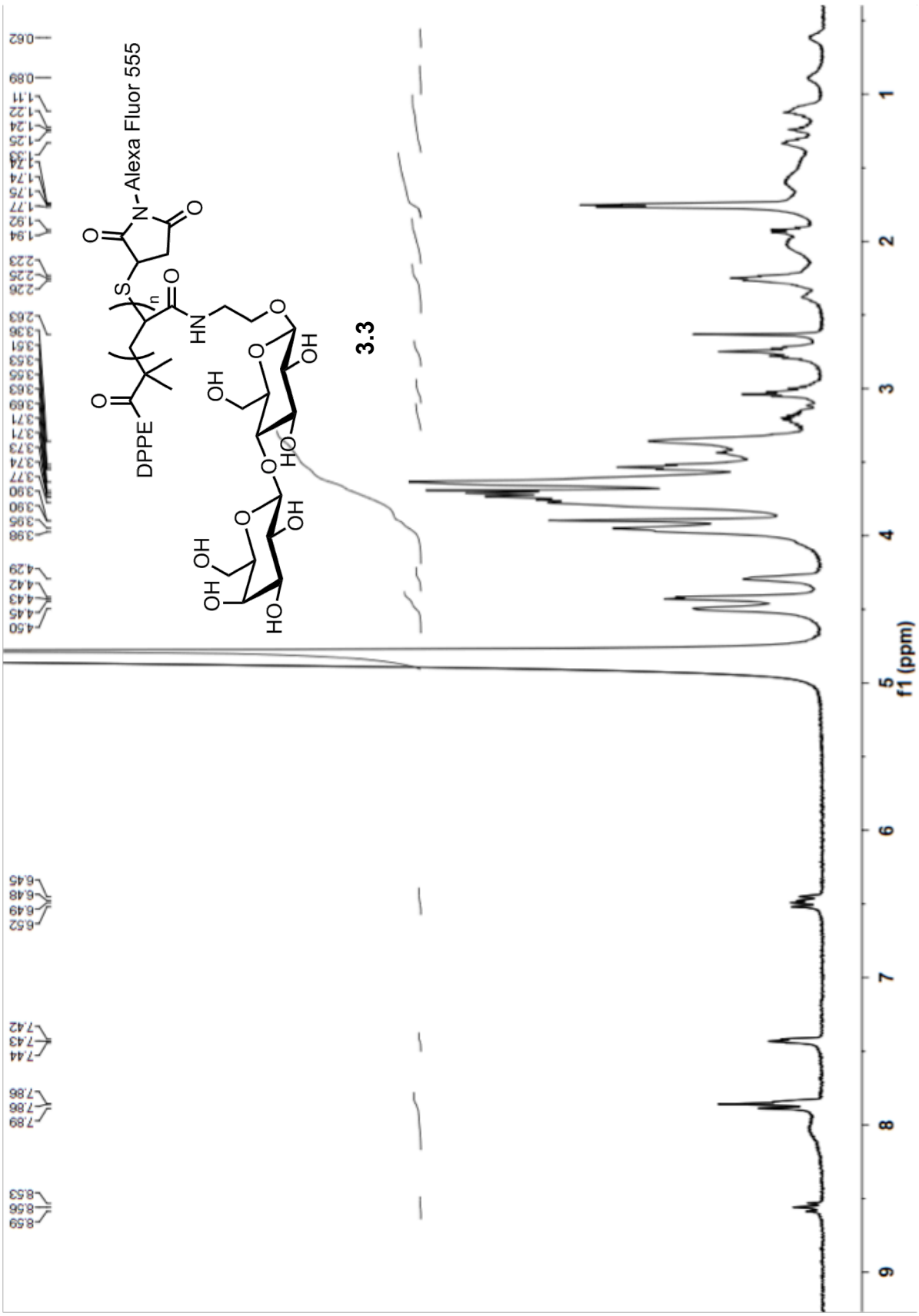

DPPE
2.24
f1 (ppm)

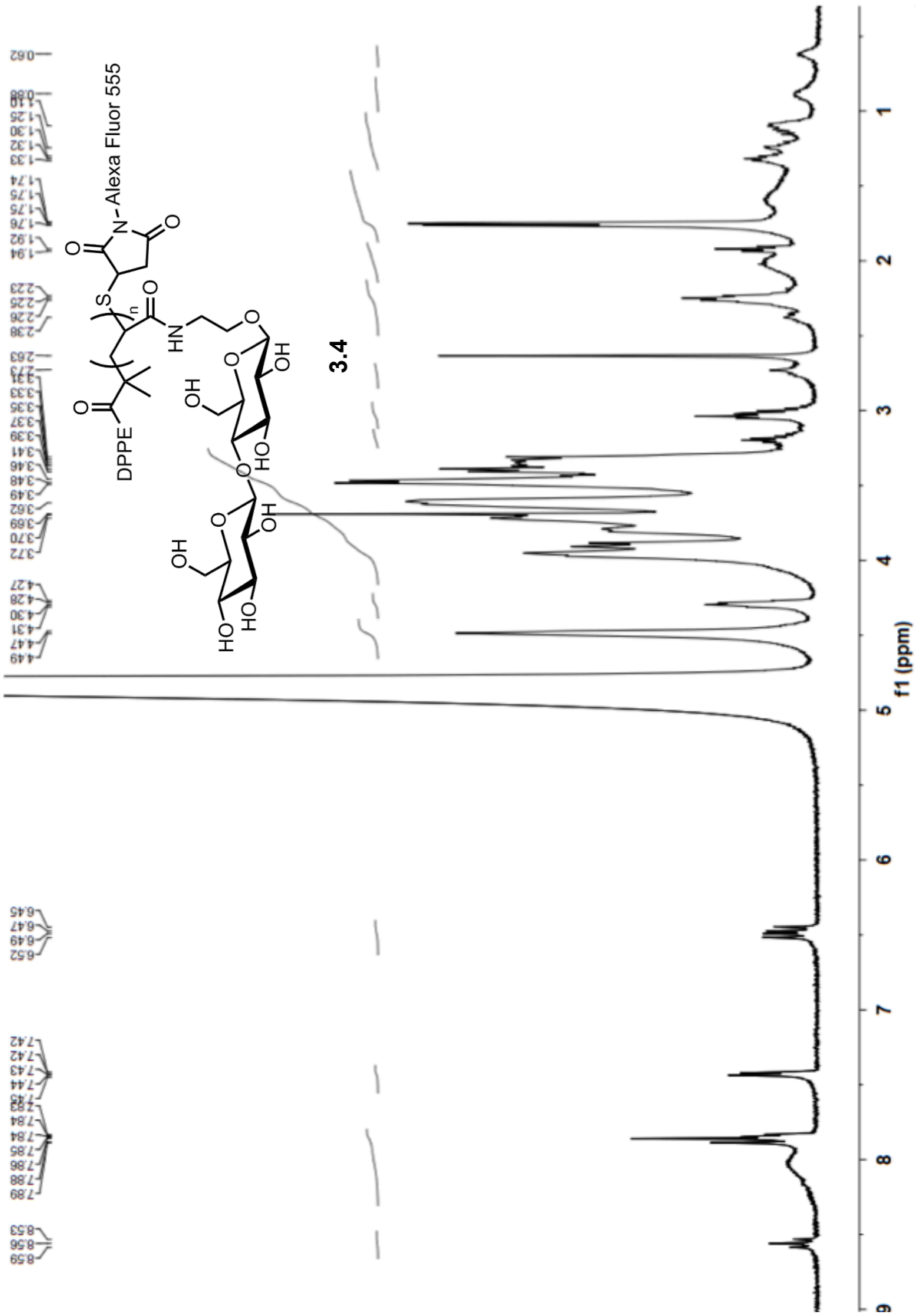

RI (au)
0
2
4
6
8
10
12
Retention Time (min)
DPPE
O
S
S
S
n
O
HN
OH
HO
HO
O
OH
OH
HO
O
O
OH

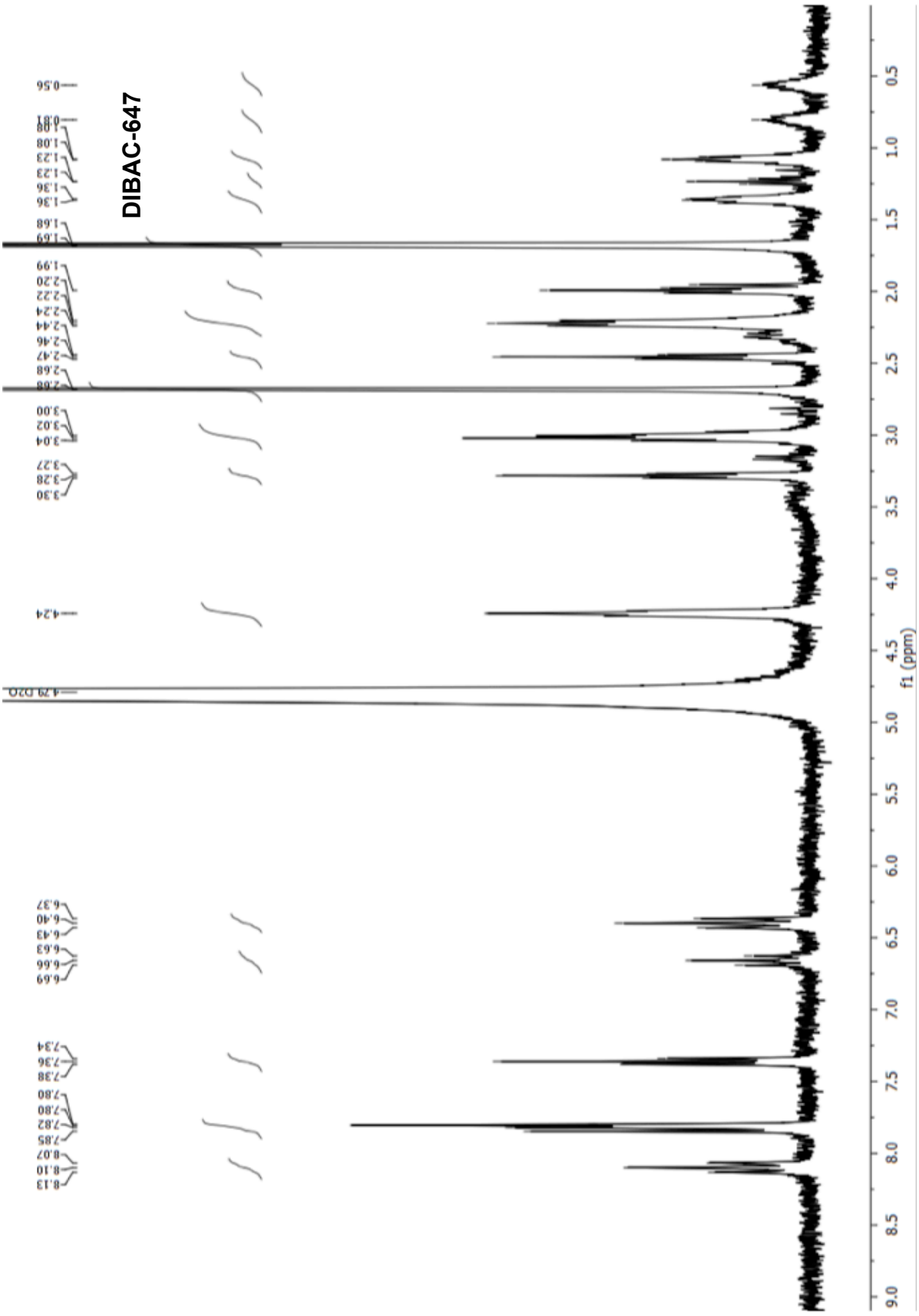
2.24

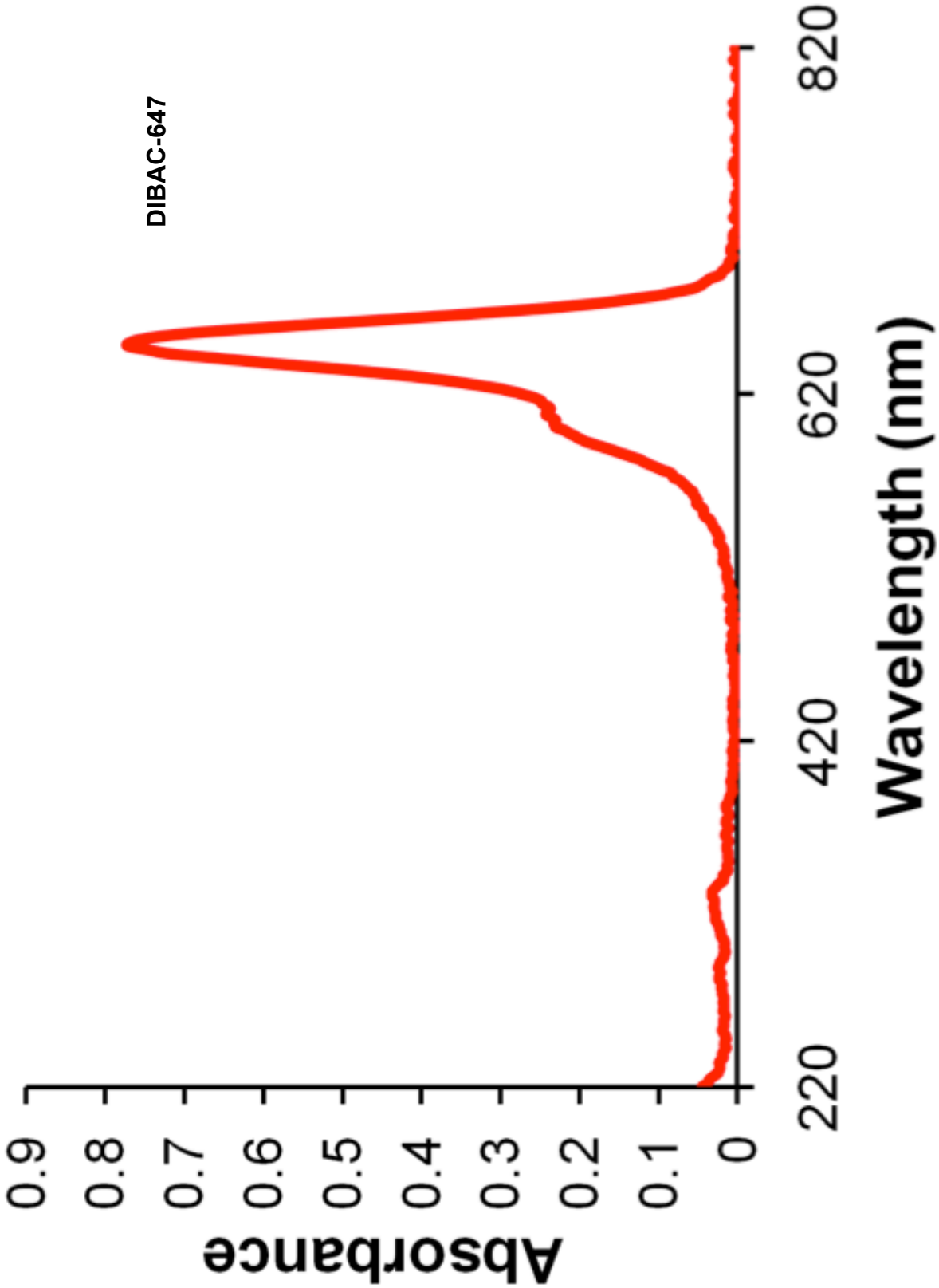

8.17
8.14
7.89
7.32
7.30
7.09
7.07
6.81
6.79
4.38
4.31
3.83
3.78
3.66
3.64
3.61
3.59
3.51
3.43
3.41
3.39
3.30
3.24
3.19
2.59
2.50
2.07
1.90
1.59
1.47
1.13
1.08
1.02
0.71
DPPE
S
n
N
Alexa Fluor 488
HN
OH
HO
3.1
9
8
7
6
5
4
3
2
1
f1 (ppm)

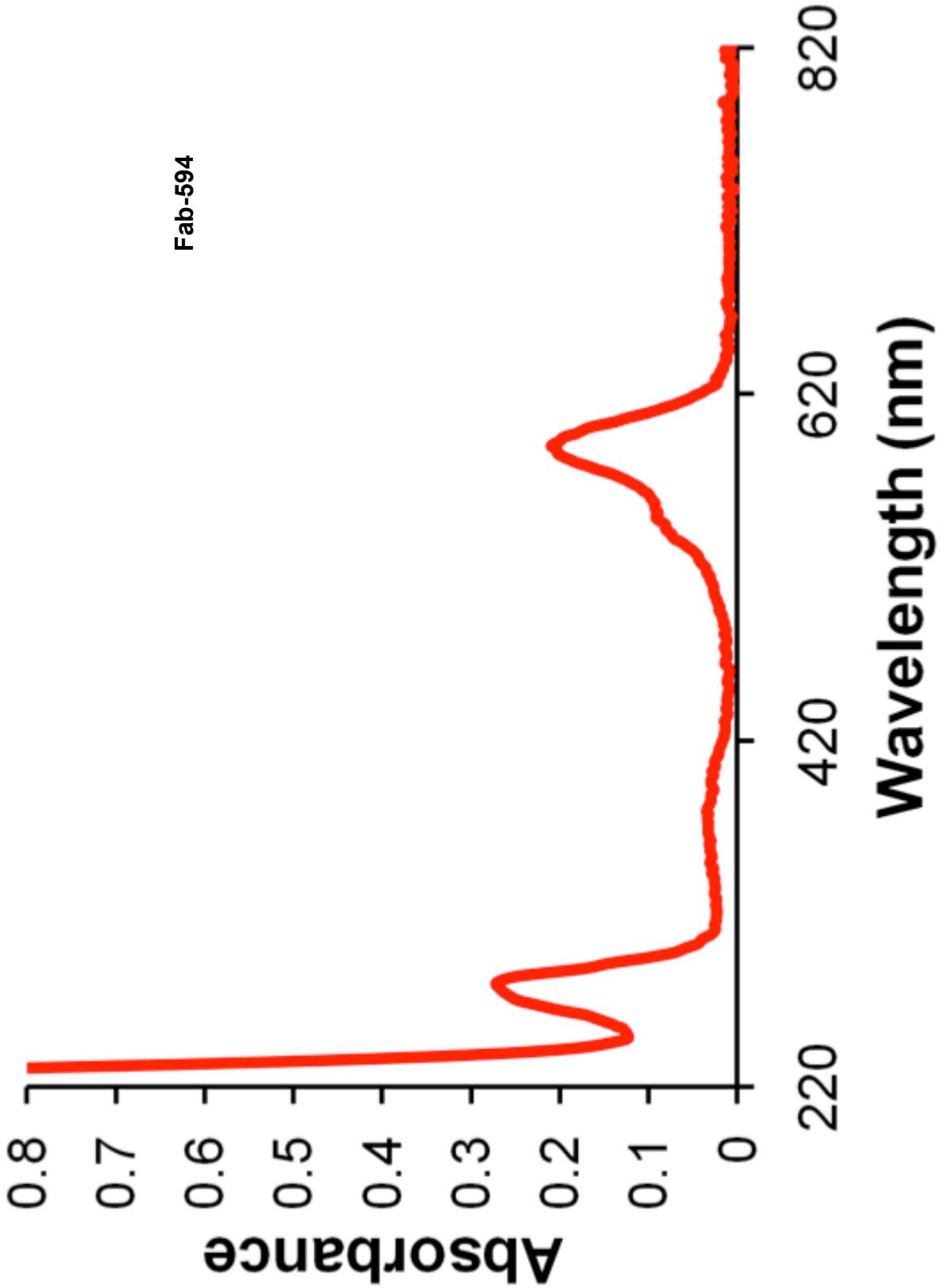

Alexa Fluor 488
DPPE
3.2
f1 (ppm)

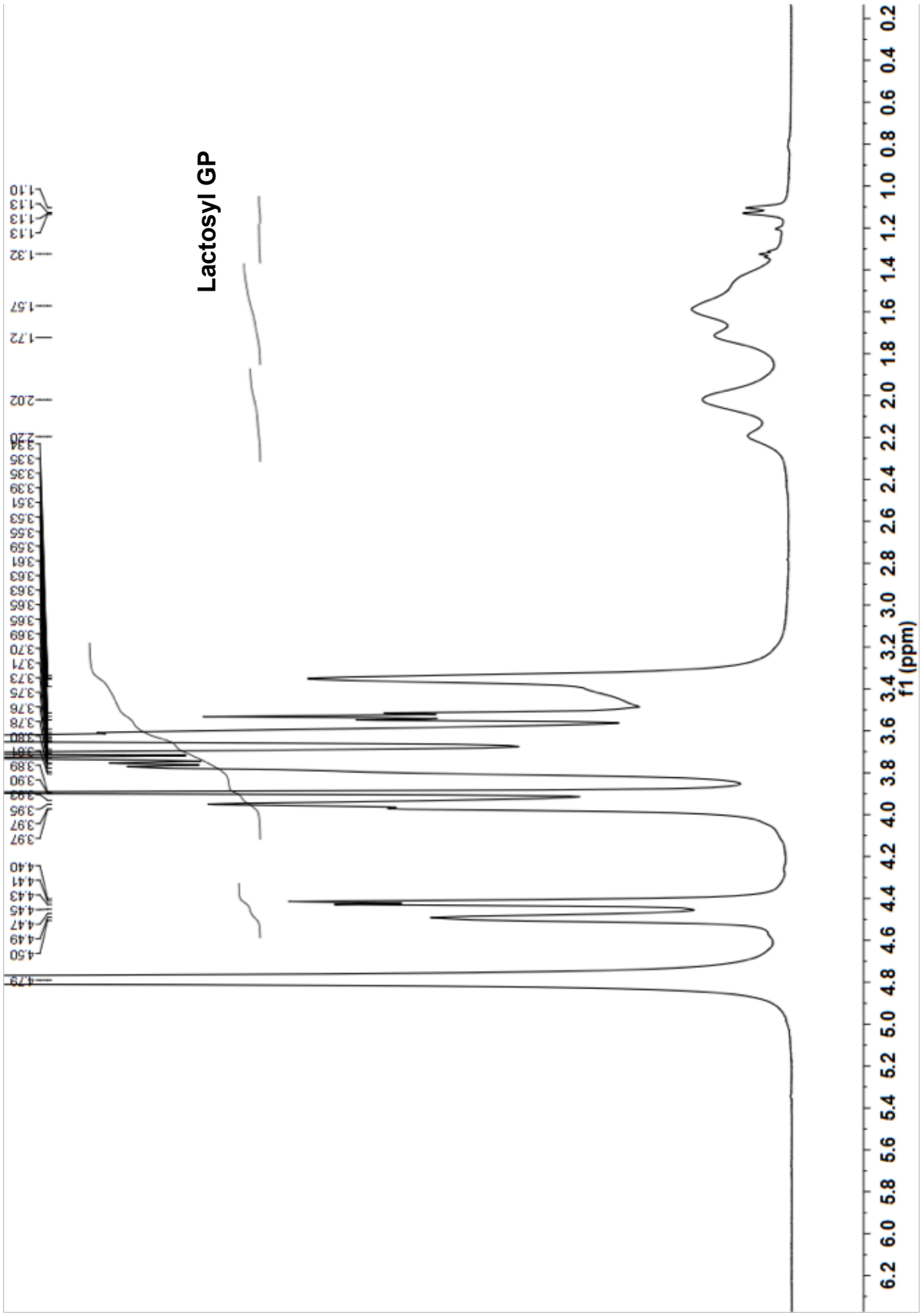

Alexa Fluor 555
DPPE
3.3
f1 (ppm)

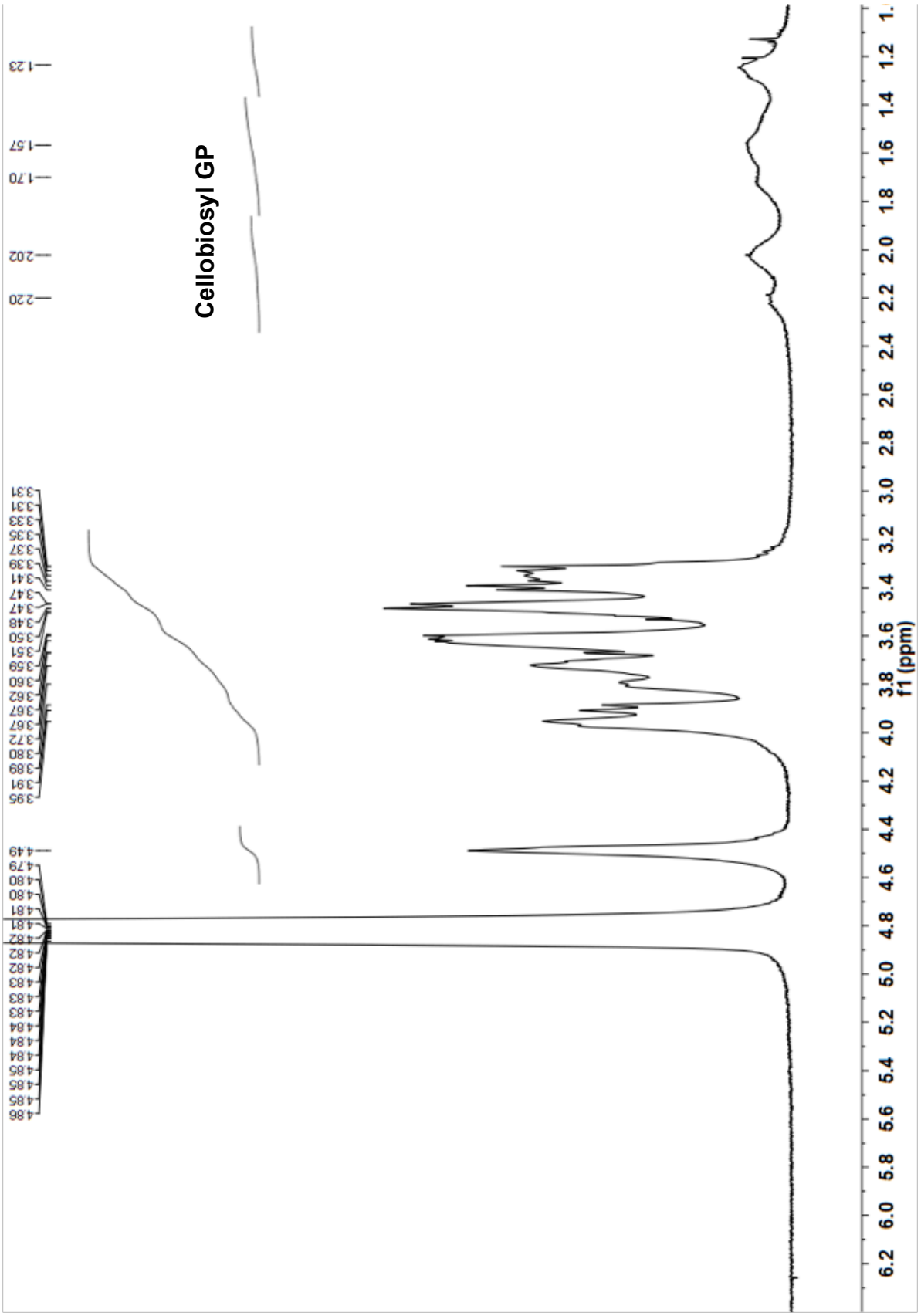

Alexa Fluor 555
DPPE
3.4
f1 (ppm)


DIBAC-647
f1 (ppm)


DIBAC-647
Absorbance
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
220
420
620
820
Wavelength (nm)


Fab-594
Absorbance
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
220
420
620
820
Wavelength (nm)


Lactosyl GP
f1 (ppm)


Cellobiosyl GP
f1 (ppm)